

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 56.4 Seconds

(without alignments)

842.900 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSPFLFQEDKR.....LSHVQDCSTPGMNSPESKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	AAW73963	Bos taurus sialida
2	1916.5	81.6	428	AAW73964	Human sialidase pr
3	1916.5	81.6	461	AAW93971	Human protein sequ
4	692	29.5	379	AAAG2547	CHO cell sialidase
5	295	12.6	245	AAU22595	Novel human colon
6	295	12.6	245	AAU22595	Human digestive sy
7	255.5	10.9	415	AAW92268	Human lysosomal si
8	213	9.1	151	AAW90545	Human immune/haema
9	192.5	8.2	492	AAW34503	Porphorymonas ging
10	192.5	8.2	540	AAW34376	Porphorymonas ging
11	176.5	7.5	516	AAU55185	Propionibacterium

12	154.5	6.6	544	17	AAW02207	Bacteroides fragil
13	146	6.2	396	15	AAW47061	Bacteroides fragil
14	140.5	6.0	683	22	AAU03632	Group B Streptococ
15	138	5.9	248	21	AAW53418	Human colon cancer
16	137.5	5.9	394	22	AAW31465	C glutamicum prote
17	136	5.8	795	22	AAU45515	Propionibacterium
18	110	4.7	51	16	AAAG2556	Tryptic peptide no
19	108.5	4.6	642	20	AAU01541	Alpha(2-3) trans-s
20	106.5	4.5	1060	20	AAU01540	Trypanosoma cruzi
21	104.5	4.5	763	22	ABG00868	Novel human diagn
22	102	4.3	680	20	AAU24320	Mouse dephosphoryl
23	99	4.2	567	22	AAW38658	Human polypeptide
24	98	4.2	931	20	AAW96253	Mouse semaphorin r
25	97	4.1	925	20	AAW96308	Neuropilin-2. Rat
26	94.5	4.0	596	22	AAW40444	Human polypeptide
27	94.5	4.0	2820	22	ABW63296	Drosophila melanog
28	94	4.0	209	22	AAW66017	Propionibacterium
29	93	4.0	244	22	AAW79141	Corynebacterium gl
30	93	4.0	244	22	AAW79142	Corynebacterium gl
31	92	3.9	479	22	AAW67522	Amino acid sequenc
32	92	3.9	567	21	AAW93988	Amino acid sequenc
33	92	3.9	628	22	AAW65805	Human leucine-rich
34	92	3.9	628	22	AAW84469	Amino acid sequenc
35	92	3.9	628	22	AAW67523	Amino acid sequenc
36	92	3.9	636	22	AAW96249	Novel human secret
37	92	3.9	909	20	AAW96249	Rat semaphorin rec
38	91.5	3.9	782	18	AAW19764	Her2-GM-CSF immuno
39	91.5	3.9	926	20	AAW96252	Mouse semaphorin r
40	90.5	3.9	529	22	AAW65308	Propionibacterium
41	90	3.8	66	22	ABB17728	Human nervous syst
42	90	3.8	380	22	ABG07024	Novel human diagn
43	89.5	3.8	1246	20	AAW06296	Human transcriptio
44	89	3.8	493	22	ABW63746	Drosophila melanog
45	89	3.8	1572	18	AAW27160	Mouse receptor ME2

ALIGNMENTS

RESULT 1

AAW73963

ID AAW73963 standard; Protein: 428 AA.

XX AC AAW73963;

XX DT 29-APR-1999 (first entry)

XX XX Bos taurus sialidase protein sequence.

XX DE Membrane-associated ganglioside sialidase; sialic acid residue removal;

XX KW Carbohydrate structure determination; cancer; gene therapy;

XX KW brain disease; cow.

XX OS Bos taurus.

XX PN WO9853052-A1.

XX XX 26-NOV-1998.

XX PF 11-MAY-1998; 98WO-JP02072.

XX PR 22-MAY-1997; 97JP-0132174.

XX PA (MIYA-) MIYAGI-KEN.

XX PI Miyagi T, Wada T, Yoshikawa Y;

XX DR WPI: 1999-095226/08.

XX DR N-PSDB; AAX01559.

XX PT Membrane-associated ganglioside sialidase - useful in the

PT investigation of carbohydrate structures and for gene therapy of

PT brain disease

Human; primer; detection; diagnosis; antisense therapy; gene therapy.
Homo sapiens.
EP1074617-A2.
07-FEB-2001.
28-JUL-2000; 2000EP-0116126.
29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-018776.
02-MAY-2000; 2000JP-0183767.
09-JUN-2000; 2000JP-0241899.
(HELI-) HELIX RES INST.
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.
Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -
Claim 8; SEQ ID 14038; 2537pp + CD ROM; English.
The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.
Sequence 461 AA;
Query Match 81.6%; Score 1916.5; DB 22; Length 461;
Best Local Similarity 82.4%; Pred. No. 2.8e-168;
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;
QY 1 MEEVTSFSSPLFQBDKRGVYTRIPALYVPPAHTFLAPAEKRSSKDEDAHLHLVLR 60
DB 34 meevttcsfnspfrqgdggyrpallyipthtflfaekrrddedahlvlr 93
QY 61 GLRTGQSVQHEPLKSLMKATLPGRHTMPCPVWERSGYVLFICVQGHVTERQIMSG 120
DB 94 glrlgqlvqgplkplmeatlpghrtmnpvcvweqsgcvflfcvgrhvtqrqivsg 153
QY 121 RNPALCLFCSQDAGYSWSVDRLTEEVIGVETHWATEAVGPGHGJQLQSGRLIIPAYA 180
DB 154 rnaaricflysdqagcswsevdteevigskelkhwatfavgpghqlqsggrlvipayt 213
QY 181 YYIPWFPCFRLPYRAPHSLMIYSDDLGTGATWHHRLIKPMVTVECEVAEVIKAGHPVL 240

DB 214 yyipwffcfqlpcktrphslmiysddlgvtwhhgrlirpmvtvecevaeavtgraghpvl 273
QY 241 YCSARTPNRHRAEALSIDHCECFQKPVLSHQLCEPHGCGSVVSECPLTIPGCGQDLAG 300
DB 274 ycsartpnrcraealstdhgegffqrlalsrqlcepphgcgsvvsfrplphrcdqss 333
QY 301 EDAPAIQOSPLLCSSVRPEAGTLESWLLYSHPNKKRRVDLGIYLNQSPLEAACWSR 360
DB 334 kdaptlqgss-pgsslrleeaagtpseawillyshptarkrvgldglynatpleaacwar 392
QY 361 PWILHCGPCGYSDLAALALENEGLFCCLFECCTKQECBOIAFRLETFDRILSHVQDCSTPG 420
DB 393 pwilhcpgcgyssdlaaleeeglfccfctkqecqiafrlthreilshlgdctspg 452
QY 421 MNSEPSK 427
DB 453 rn--psg 457
RESULT 4
AAR62547
ID AAR62547 standard; Protein; 379 AA.
AC AAR62547;
XX
DT 09-AUG-1995 (first entry)
XX
DE CHO cell sialidase.
XX
KW Sialidase; sialic acid; Chinese Hamster ovary.
OS Crictetus cricetus.
FH Key Location/Qualifiers
FT Peptide 11..29
FT Peptide 112..136
FT Peptide 45..75 /label= PCR 14/17 probe
FT Peptide 139..169
FT Peptide 172..201
FT Peptide 208..218
FT Peptide 244..263
FT Peptide 284..335
FT Peptide 231..237
PN WO9426908-A.
XX
PD 24-NOV-1994.
XX
PF 17-MAY-1994; 94WO-US05471.
XX
PR 17-MAY-1993; 93US-0062586.
PR 25-JAN-1994; 94US-0187327.
XX
PA (GETH) GENENTECH INC.
XX
PI Sliwkowski MB, Warner TG;
XX
DR WPI; 1995-006802/01.
DR N-PSDB; AAQ75442.
XX
PT Recombinant cell line defective for sialidase expression - useful for prodn. of glyco:protein(s), also new CHO cell sialidase and related DNA etc., for treating inflammation, lung diseases, etc.
PS Disclosure; Figure 10; 64pp; English.
XX
CC Purified sialidase from CHO cells was treated with various proteases and the resulting peptides sequenced (see, for example, AAR62565). The lysine C derived peptide LC18 (see AAR62565) is a composite of tryptic peptides TP14 and TP17 and is referred to as TP14/17. A PCR 14/17 probe was prepd. and used

CC as a probe. A Chinese hamster ovary cell - lambda
CC gt 10 cDNA library (JL 1001a, obtd. from Clontech Labs, CA)
CC was screened using radiolabeled PCR 14/17 probe. One plaque
CC referred to as clone 15, gave a strong hybridisation signal.
CC The DNA sequence of clone 15 was determined to be AAQ75442 and
CC the predicted AA sequence to be AAR62547. The posns of the protease
CC derived peptide fragments and the sequence corresp. to the PCR
CC 14/17 probe are indicated in FT.
XX
SQ Sequence 379 AA;

Query Match 29.5%; Score 692; DB 16; Length 379;
Best Local Similarity 38.6%; Pred. No. 2.2e-55;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;

QY 13 LFQOEDKRGVTRIPALIVYPPAHTFLAPAEKRSSKDEDLHLVLRRLGTGQS--VQW 70
DB 12 lfqtgd--yayripaliylskqktllafaeakrltktddehadvlrrgsgynadthqvw 68
QY 71 EPLKSLMKATLPGHRTWPCVWPKRGVYVLFVFCVQGHVTERQOIMSGRNPRLCFTC 130
DB 69 qaeenvtqayleghrmspcplydkqtrtlfiffiavrgqisehqlgtgvnvtlchit 128
QY 131 SODAGYSWSDVRLTEEVIGPEVTHWATFVAGVPGHGILQIQ--SGRLIIPAYAYIIPFWFF 188
DB 129 stdhgktwsavqldtdttgstgtdwatfgvpgphclqrlntagsllvpayayr----- 182
QY 189 CTRFLP--YRAPHSLMIYSDDGATWHHGLRIKPMVTVECEVAEIVGKAGHPVLYCSART 246
DB 183 --kqpplhapapsafclshdngstwelghfvs-qnslecqvaev-gtgaervvylinars 238
QY 247 PNRHRAEALSIDHGECFORPVLSHOLCEPPHCGCQSVSFCPLPIPGQCQDLAGEDAPAI 306
DB 239 clgarvqadspusgidfqdnqvsklveppkgchgsviaf----- 278
QY 307 QOSSLPCSSVRPEPEAGTILSESLLYSHPTNKKRRVDLGIYLNQSPLEAACWSRPWLHC 366
DB 279 -----pnptskadal-dwlllythptdsarkrtngvynqkpldpttwsaptllat 328
QY 367 GPCGYSDLAAL-----ENEGFLCFLFCGCKQCEQIAFLRFTDREILSHVQ 414
DB 329 gicaysdlqnmhgdpqspqfgyclve---snnyeelvfmlftlkqafpavfg 377

RESULT 5
AAU22595
ID AAU22595 standard; Protein; 245 AA.
XX
AC AAU22595;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human colon associated polypeptide #128.
XX
KW Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW cytostatic.
OS Homo sapiens.
XX
PN WO200155302-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01240.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 28-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0214886.
PR 07-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Harash SC, Ruben SM;
XX
XX WPI: 2001-502630/55.
XX N-PSDB; AAK88342.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases -
XX
XX Claim 11; SEQ ID NO 1918; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a digestive system antigen of
XX the invention.
XX
XX Sequence 245 AA;

Query Match 12.6%; Score 295; DB 22; Length 245;
Best Local Similarity 62.1%; Pred. No. 5.4e-19;
Matches 54; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

```
QY 97 SGVYVLFICVGHVTERQIIMSRNPRLCFICSDAGYSWSDVRDLTEEVIGPEVTHW 156
ID AAW69228 standard; Protein; 415 AA.
Db 5 tgvclfflavghtpeavqiaqgrnaarccvasrdaglsqgsardlteeaiggaqvqw 64
QY 157 ATFVAGPGHGICQISQRLIIPAYAYI 183
Db 65 atfavpgghgvqlpsgrllvpxytrv 91

RESULT 7
AAW69228
ID AAW69228 standard; Protein; 415 AA.
AC AAW69228;
XX
XX
DT 18-FEB-1999 (first entry)
DE Human lysosomal sialidase.
KW Lysosomal sialidase; human; sialidosis; lysosomal storage disease;
KW Sandhof disease; mutation detection; Tay-Sachs disease; therapy.
XX
OS Homo sapiens.
XX
PN WO9831817-A2.
XX
PD 23-JUL-1998.
XX
PF 13-JAN-1998; 98WO-CA00026.
XX
PR 14-JAN-1997; 97US-0035092.
XX
PA (HOPI-) HOPITAL SAINTE-JUSTINE.
XX
PI Potier M, Pshezhetsky AV;
XX
DR WPI; 1998-414113/35.
XX
DR N-PSDB; AAV44790.
XX
XX
PT New human lysosomal sialidase and related nucleic acid - used to
PT detect mutation(s) that cause sialidosis and similar disease, also
PT for treating these diseases and screening for antiviral agents
PS Claim 1; Fig 1; 30pp; English.
XX
CC This sequence is the human lysosomal sialidase of the invention. The
CC DNA is used as a reference to identify mutations that cause sialidosis or
CC similar disease and for chromosome mapping. The protein is used:
CC (i) for treating lysosomal storage diseases (sialidosis, Tay-Sachs
CC disease or Sandhof disease); (ii) to screen for agents useful against
CC viral sialidase (potentially useful as antiviral agents without side
CC effects on the human enzyme); (iii) for digesting sialylated
CC oligosaccharides or glycolipids in milk; and (iv) when inactivated, as
CC antiviral agents (by binding to surface sialic acid with high affinity,
CC so preventing binding of virus to cells).
XX
SQ Sequence 415 AA;

Query Match 10.9%; Score 255.5; DB 19; Length 415;
Best Local Similarity 28.0%; Pred. No. 4.9e-15;
Matches 116; Conservative 46; Mismatches 130; Indels 123; Gaps 23;

QY 23 TYRIPALIVPPANTFLAFAEKRS-SSKDEDAHLVLRGLRTGQSVQWPEPLKSLMK-AT 80
Db 76 tfrilp-litatrptlatafearkmssdegakfiarlrsmddgst--wsptafivndgd 132
QY 81 LPGHRTMNPCEWVERSGYVYLPF-TCVQGHVTERQIIMSRNPRLCFICSDAGYSWS 139
Db 133 vpdglnlg-avvadvetgvvflfysicah-----kagcqvasmtlwwskddgvs 182
QY 140 DVRLDLTEEVIGPEVTHWATFVAGPGHGICQI-----SGRLIIPAYAYIIPWFCFLPYR 195
```

```
Db 183 tprnlsld-igtev-----fapggsgikqqrtrkgrlivcghgtlerdgvfc----- 230
QY 196 ARPHSLMIYSDDLGATWHHGRLIKPMVTVECEVAEYIGRAGHPVLVYCSARTPNRHRAEAL 255
Db 231 -----llsddhgaswrygs-----gvsqip----- 250
QY 256 SIDHGECFOKPVLSHQLCPEPHGCGSVV-----SFCPLIIPGCGODLAGEDAPAIQ 307
Db 251 ----yqpkqendfnpdecqpyelpdgsvvinnarnqnnyhchriv-----lrsyda---- 298
QY 308 QSPLLCSSVRP-----EPE-----AGTL--SESLYLXSHPTNKKRRVDLGIYLNQSPL 353
Db 299 -----cdtllrpdvtdfdpelvdpvvaagavvtssgivffsnpahpefrvnl--tlrwsfs 351
QY 354 EAACWSRPWI-LHCPCGYSGLAALNEGLFG-----CLFECGTKECEBOIA 399
Db 352 ngtswrketvqlwpgpsgysslatleg-amsdgeeqapqlvlyekgrnhythesis 405

RESULT 8
AAW90525 standard; Protein; 151 AA.
XX
AC AAW90525;
XX
DT 07-NOV-2001 (first entry)
DE Human Immune/haematopoietic antigen SEQ ID NO:18118.
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189674.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225277.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
```

PR	22-AUG-2000;	2000US-02258665.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227182.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0233397.
PR	14-SEP-2000;	2000US-0233398.
PR	14-SEP-2000;	2000US-0233399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0235802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0245532.
PR	08-NOV-2000;	2000US-0245523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.

CC especially gingivitis.

XX Sequence 540 AA;

Query Match 8.2%; Score 192.5; DB 20; Length 540;

Best Local Similarity 22.7%; Pred. No. 4.6e-09;

Matches 93; Conservative 60; Mismatches 164; Indels 93; Gaps 16;

QY 13 LFOQEDKGVYRIPALIVPPAHTFLFAFAEKSSSKDE--DALHLVLRGLRTGQSQVOW 70
DQ 196 lfvpdggssrnyripailktang-tliamadrkynqtdlpedidivmrtdgks--w 252
QY 71 EPLSLMKATLPGHRTMPCPVWKRKSGYVLFVFCVGOHVTERQQIMSGRNPALCLFIC 130
DQ 253 sdprliivgegrnh-gfgdvalvtqagklmifv---ggvlgws--tpdrprtysie 306
QY 131 SODAGYSWSDVRLTEEVIG-----PEVTHWATFVGGHGIQLOSGRLIIPA-----Y 179
DQ 307 srdegltwpprdithfigkdcadpgrsrwlasfcasggglvlpstgrttfvaalresgq 366
QY 180 AYIIPFWFCFRLPYRAPHSLMIYSDDLGGATW-----HHGRLIKPMVTVECEVAEVI 232
DQ 367 eyvl-----nnyvlssddgdwtqlsdcayrgrdeaklsmpdgrvlmsl 411
QY 233 GKAGHPVLYCSARTPNRHRABALSDHGECEQKPVLSHQLCEPFGCGQSVVFCFLP 292
DQ 412 rnqg-----rqesqrffalssddglwtweraqkfegindh--gcngam-----lqvk 456
QY 293 GGCQDLAGEDAPAIQOSSLCSVRPEPAGTLESWLLYSHPTNKKRRVDLGIYLNOSP 352
DQ 457 rngrd-----qvlhslp1gpdgrrdgavylfdh- 484
QY 353 LEACWSRPWILHCGPCGYSDLAALENGLFCGLFCEGCKOCEQIAFRL 402
DQ 485 -vsgrwsapvvvngssaysdmtila-dgtigyfveesdeislvirfvl 532

RESULT 11

AAU55185

ID AAU55185 standard; Protein; 516 AA.

AC AAU55185;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #16081.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS959568.

XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

PS Claim 3; SEQ ID No 16380; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 516 AA;

Query Match 7.5%; Score 176.5; DB 22; Length 516;

Best Local Similarity 24.1%; Pred. No. 1.3e-07;

Matches 95; Conservative 52; Mismatches 158; Indels 89; Gaps 19;

QY 17 EDKRGVYRIPALIVPPAHTFLFAFAEKSSSKD-EDALHLVLRGLRTGQSQVQWEPLKS 75
DQ 127 ehpegvfyrpalatasngdlasydlrpsagdapnpsivqrrsrdngrt--wgp-qt 183
QY 76 LMKATLPGHRTM---NCPVWKRKSGYVLFPI--CVGOHVTERQQIMSGRNPALCLFI- 129
DQ 184 vihaqtgprkrkvysdpsylvdpatgrilnfhvksydrfatse-----vgtdpdrhrvlh 239
QY 130 ---CSQDAGYSWSDVRLTEEVIGPEVTHWATFVGGHGIQL-----QSGRLIIPAYAY 181
DQ 240 aevststdnghtwth-rditreitsdpttr--trfvasggqiallhghagrliag---- 292
QY 182 YIPWFVFCFRLPYRAPHSLMIYSDDLGGATWHHGLIKPMVTVECEVAEVIKAGHPVLY 241
DQ 293 -----mtvrnsvggqagsyddhgitwhagnpvgmmnd-enkvvel-----sdgtlm 339
QY 242 CSARTPNR--HRABALSDHGECEQKPVLSHQLCEPFGCGQSVVFCFLPFGCQDLA 299
DQ 340 lnsrdaarsgrkvaysqdggltwgpkvlvddldtptnna----- 380
QY 300 GEDAPAIQOSSLCSVRPEPAGTLESWLLYSHPTNKKRRVDLGIYLNOSPLEAAC-W 358
DQ 381 -----liraypnaragsakarillftnarnatervn-----gtlsvscdd 420
QY 359 SRPWILH---CGPCGYSDLAALENGLFCGLFE 388
DQ 421 grtwshqtympegvgyt-taavsdgdlgvlwe 453

RESULT 12

AAW02207

ID AAW02207 standard; Protein; 544 AA.

XX AAW02207;

XX 11-OCT-1996 (first entry)

XX Bacteroides fragilis neuraminidase.

XX

KW Neureminidase; identification; diagnosis; pathogenic factor;
 KW stabilisation; sialic acid containing substance.

OS Bacteroides fragilis.

PN JP08131174-A.

XX 28-MAY-1996.

XX 15-NOV-1994; 94JP-0280237.

XX 15-NOV-1994; 94JP-0280237.

XX (SAKA) OTSUKA PHARM CO LTD.

XX WPI; 1996-303851/31.

DR N-PSDB; AAT36720.

XX Bacteroides fragilis neureminidase gene - useful for identification
 PT and diagnosis of B. fragilis

XX Claim 2; Pages 9-11; 12pp; Japanese.

CC The present sequence is the B. fragilis neureminidase (NA),
 CC which was extracted from B. fragilis using, e.g. the method of
 CC Thomas (Berns, L.I. and Thomas, Jr., C.A. J. Mol. Biol., 11,
 CC 476-490 (1965)). The whole or a part (pref. an EcoRI-Cla fragment)
 CC of the NA gene can be used for the identification and diagnosis of
 CC B. fragilis, a microorganism believed to have a NA a pathogenic
 CC factor not considered to be present in other bacterial spp.. The
 CC NA may be used to stabilise a sialic acid-contg. substance.

XX Sequence 544 AA;

Query Match 6.6%; Score 154.5; DB 17; Length 544;
 Best Local Similarity 21.6%; Pred. No. 1.5e-05;
 Matches 84; Conservative 59; Mismatches 157; Indels 89; Gaps 15;

QY 24 YRIPALIVPPAHTFLAFAEKRSKSD-EDALHLVLRGLRTGQSVQWE-----PLKSLMK 78

DB 203 frlpplvttnkgtllgvdyvrynsdvldqehvdlgrstddgkt--wekmrlplafgef 260

QY 79 ATPLPG--HRTMNPVPCVVERKSGYVLFVFCVQGHVTER-----QQIMSGRNPALCFICS 131

DB 261 gglpaggngvgdpsilvdtktnnvvvaawthgmgngqrawsshpqmdmhtaglvlaaks 320

QY 132 QDAGYSNDVRDLTEEVIGPEVTHMATFVAVGPHGHIQLOSGRLIIPAYAYIIPFWFFCFR 191

DB 321 tddgktsapiniteqvkdps---wyflgpggrgltmsdgtlvftqf----- 366

QY 192 LPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVIKAGHPVLVCSARTPNR-- 249

DB 367 idstrvnpagimyskdggnkwmhnyartn--tteaqvaev-----epgvlmnmrdnrg 420

QY 250 HRBAEALSDHGEFCQKPVLSHQLCEPPHGCQGSVVSPCLPIPGCCQDLAGEDAPAIQOS 309

DB 421 sravaiktdlgktwtehesark-----alpes 447

QY 310 PLLCSSLVRPEPEAGTLESWLLYSHTNKKRRVDLGIYLNQSPLEAACWS----- 359

DB 448 vcmassliisvakdnvlgkdlifsnpttkgrynttikisid--ggvtspehqlildeg 505

QY 360 RPWLHCGPCGYSDLAALENEGLFGCLFE 388

DB 506 nnw-----gysclsmidketi-gliye 526

RESULT 13

AAR47061

ID AAR47061 standard; Protein: 396 AA.

XX

AC AAR47061;

XX 03-OCT-1994 (first entry)

XX Bacteroides fragilis neureminidase.

XX Neureminidase; B. fragilis; detection.

XX Bacteroides fragilis.

XX JP06046865-A.

XX 22-FEB-1994.

XX 17-JUL-1991; 91JP-0269931.

XX 17-JUL-1991; 91JP-0269931.

XX (SAKA) OTSUKA PHARM CO LTD.

XX WPI; 1994-097024/12.

XX N-PSDB; AAQ58660.

XX Neureminidase gene from Bacteroides fragilis - useful for prodn.
 PT of neureminidase and detection of B. fragilis

XX Claim 2; Page 10-11; 14pp; Japanese.

XX The gene from B. fragilis YCH46 coding for neureminidase AAR47061 is
 CC useful for recombinant production of the enzyme and for the
 CC identification and diagnosis of B. fragilis infection.

XX Sequence 396 AA;

Query Match 6.2%; Score 146; DB 15; Length 396;
 Best Local Similarity 21.5%; Pred. No. 5.9e-05;
 Matches 84; Conservative 59; Mismatches 157; Indels 90; Gaps 16;

QY 24 YRIPALIVPPAHTFLAFAEKRSKSD-EDALHLVLRGLRTGQSVQWE-----PLKSLMK 78

DB 54 frlpplvttnkgtllgvdyvrynsdvldqehvdlgrstddgkt--wekmrlplafgef 111

QY 79 ATPLPG--HRTMNPVPCVVERKSGYVLFVFCVQGHVTER-----QQIMSGRNPALCFICS 131

DB 112 gglpaggngvgdpsilvdtktnnvvvaawthgmgngqrawsshpqmdmhtaglvlaaks 171

QY 132 QDAGYSNDVRDLTEEVIGPEVTHMATFVAVGPHGHIQLOSGRLIIPAYAYIIPFWFFCFR 190

DB 172 tddgktsapiniteqvkdps---wyflgpggrgltmsdgtlvftqf----- 218

QY 191 LPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVECEVAEVIKAGHPVLVCSARTPNR-- 249

DB 219 -idstrvnpagimyskdggnkwmhnyartn--tteaqvaev-----epgvlmnmrdnrg 271

QY 250 -HRAEALSDHGEFCQKPVLSHQLCEPPHGCQGSVVSPCLPIPGCCQDLAGEDAPAIQO 308

DB 272 sravaiktdlgktwtehesark-----alpe 298

QY 309 SPLCSSLVRPEPEAGTLESWLLYSHTNKKRRVDLGIYLNQSPLEAACWS----- 359

DB 299 svcmassliisvakdnvlgkdlifsnpttkgrynttikisid--ggvtspehqlilde 356

QY 360 -RPWLHCGPCGYSDLAALENEGLFGCLFE 388

DB 357 gnnw-----gysclsmidketi-gliye 378

RESULT 14

AAU03632

ID AAU03632 standard; Protein: 683 AA.

XX

AC AAU03632;

Db 66 sddhgaswrygs-----gvsgip-----ygpqk 89
Qy 265 KPVLSHQLCPEPHGCQGSV-----SFCPLPCCQDLAGEDAPAIQQSPLLCSSV 316
Db 90 endfnpdecqpyelpdgsvvlnarngnnyhchriv-----lrsyda-----cdtl 135
Qy 317 RP-----EPE-----AGTL--SESWLLYSHTNKKRRVDLGIYLNOSPLEAACWSRPW 362
Db 136 rprdvtfdpelvdpvvaagavvtssglvffsnpahpefrvnl--tlrwsfngtswrket 193
Qy 363 I-LHCGPCGYSDLAALENEGLFG-----CLFECGTKQCEQIA 399
Db 194 vqlwpgpsgysslatleg-smdgeeapqlylyvekgrnhythesis 238

Search completed: October 7, 2002, 15:00:52
Job time: 134 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 32.29 Seconds
(without alignments)
1273.652 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSSFLFOQEDKR.....LSHVQDGDSTPGMNSPPSKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1518	64.7	418	2 JC7588	exo-alpha-sialidas
2	708.5	30.2	385	2 JC7751	exo-alpha-sialidas
3	705	30.0	379	2 A49679	exo-alpha-sialidas
4	692	29.5	379	2 A54961	sialidase, cytosol
5	259	11.0	648	2 T34999	probable neuramida
6	201.5	8.6	647	2 A45244	exo-alpha-sialidas
7	197	8.4	748	2 T37097	probable secreted
8	186	7.9	1014	1 NMCLSS	exo-alpha-sialidas
9	172	7.3	694	2 I40866	exo-alpha-sialidas
10	169	7.2	901	2 A49227	sialidase - Actino
11	154	6.6	913	2 S20590	exo-alpha-sialidas
12	149	6.3	1035	2 T30287	exo-alpha-sialidas
13	149	6.3	1035	2 G38063	exo-alpha-sialidas
14	146	6.2	545	2 JC2500	exo-alpha-sialidas
15	145.5	6.2	404	2 A37234	exo-alpha-sialidas
16	128	5.5	382	2 S01339	exo-alpha-sialidas
17	125	5.3	740	2 G35153	neuraminidase, pro
18	120	5.1	382	2 S32148	exo-alpha-sialidas
19	117.5	5.0	773	2 JF0387	exo-alpha-sialidas
20	107	4.6	1548	2 S34583	serine proteinase
21	106.5	4.5	879	2 S23006	shed acute-phase a
22	105	4.5	781	2 A43866	neuraminidase - VI
23	105	4.5	807	2 E82158	neuraminidase VC17
24	102	4.3	680	2 JC5895	killer cell inhibi
25	101	4.3	376	1 NNEBST	exo-alpha-sialidas
26	100	4.3	2767	1 UTHU	thyroglobulin prec
27	95.5	4.1	389	2 T08135	sedoheptulose-bisp
28	94.5	4.0	1216	2 A55620	apical endosomal p
29	91	3.9	697	2 E95196	neuraminidase B [1

RESULT 1

JC7588

exo-alpha-sialidase (EC 3.2.1.18) - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C:Accession: JC7588

R:Hasegawa, T.; Carnero, C.F.; Wada, T.; Itoyama, Y.; Miyagi, T.

Biochem. Biophys. Res. Commun. 280, 726-732, 2001

A:Title: Differential expression of three sialidase genes in rat development.

A:Reference number: JC7588; MUID:21092669; PMID:11162581

A:Contents: Brain

A:Accession: JC7588

A:Molecule type: mRNA

A:Residues: 1-418 <HAS>

A:Cross-references: DDBJ:AB026841

C:Comment: This enzyme is involved in metabolic regulation and cellular functions, in

C:Keywords: glycosidase; hydrolase

F:24-27/Region: consensus motif YRIP

F:131-138/Region: consensus Asp-box, three copies #status predicted

F:173-195/Domain: hydrophobic #status predicted <HYD>

F:203-210/Region: consensus Asp-box, three copies #status predicted

F:254-261/Region: consensus Asp-box, three copies #status predicted

Query Match 64.7%; Score 1518; DB 2; Length 418;

Best Local Similarity 65.9%; Pred. No. 7e-123;

Matches 278; Conservative 58; Mismatches 82; Indels 4; Gaps 2;

QY 1 MEEVTSFSSFLFOQEDKRGVYTRIPALYYPPAHTFLAFAEKRSSKDEALHLVLR 60

DB 1 MEEVSSCLSRSLFQOEQNEITYRIPALLYIPPTHTFLAFAEKRTSSHDEAVYLVFRR 60

QY 61 GLRTQSQVQWEPLKSLMKATLPGHRTMNPQVWERKSGYVYLFVFCVQGVHTERQIMSG 120

DB 61 GVMKCSVEMGQQQLMEATLPGHRTMNPQVWERKSGYVYLFVFCVQGVHTERQIMSG 120

QY 121 RNPARLCFTCSODAGYSWSDVRDLTEEVIGPEVTHATFVAGPGHGIIQSGRLIIPAYA 180

DB 121 RNARALCELYSDSCSGEVDKLTVEEVVSGEMKHWATFVAGPGHGIIQSGRLIIPAYA 180

QY 181 YIYPMFCCFRLPYRARPISLMYSDDLGATWHGRLTKPMYTVCEVAEVIKAGHPVL 240

DB 181 YLISCFWLCF--PCSVKPHSLMFYSDDLGVTHWCGRFKIPQVTGECQVAEVPKAGNLVL 238

QY 241 YCSARTPNRHRAEALSIDHGECFQKPVLSHQICEPHGCGQSVSWFCEIPGCGQDLAG 300

DB 239 YCSARTPNKFAEAFTSDGDCFQKPVLSHQICEPHGCGQSVSWFCEIPGCGQDLAG 300

QY 301 EDAPAIQSPPLICSSVRPEPEAGTLESWLLYSHPTNKKRRVDLIYILNQSPLEACNWR 360

DB 299 KDVPSTQKPLMDRSLEVEEGAGPSTGLWLYSHPTNKKRNLGLIYINQNPLEVYNWR 358

hypothetical prote
surface antigen gp
pyruvate flavodoxi
conserved hypothet
Photinus-luciferin
Photinus-luciferin
BFLF1 protein - hu
phage-related prot
killer cell inhibi
exo-alpha-sialidas
hypothetical prote
hypothetical prote
hypothetical prote
carboxypeptidase-r
C4BP alpha chain p

A:Cross-references: GB:U06143; NID:g509824; PIDN:AAA19746.1; PID:g509825

Query Match 29.5% Score 692; DB 2; Length 379;
Best Local Similarity 38.6% Pred. No. 8.3e-52;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;
Qy 13 LFOQEDKRGVYRIPALYVPPAHTFLAFABKRSKDEDAHLVLRGLRTGOS--VQW 70
Db 12 LFQTDG---YAIRIPALYLSKQTLTLLAFABKRLTKTDEHADLFVLRGSSYNADTHQVQW 68
Qy 71 EPLKSLMKATLPGHRTMPCPWERKSGYVYLFFICVGHVTERQOIMSGRNPRLCEPIC 130
Db 69 QAEEVVTQAYLEGRHSMSPCLYDKQRTLEFLFTAVRGQISEHHQLQTGVNVTRLCHIT 128
Qy 131 SODAGYSWDRDLTEEVIGPEVTHWATFVAGPGHGIQLO--SGRLIIPAYAYIPIFWFF 188
Db 129 STDHGKWSAVQDLDTTTIGTHQDMATFVGPGHCLQRLNRTAGSLVPAFYR----- 182
Qy 189 CRLP--YRARPRLSMYSDDLGATWHHGRILKPMVTVECEVAEYIGKAGHPVLVYCSART 246
Db 183 --KOPPIHAPAPSAFCFLSHDHGHWELGHEVS-QNSLECOVAEV-GTGAERVVYLNARS 238
Qy 247 PNHRRAEALSDHGECEFOKPVLSHQLECEPHGCGSVVFCPLIPGCGODLAGEDAPAI 306
Db 239 CLGARVQAQSPNSGLDFODNQWSKLVBPCKGCHGSVIAF----- 278
Qy 307 QOSPLLCSSVPEPEAGTSLSEWLLYSHPTNKKRRVDLGIYLNOSPLEAACWSRPIILHC 366
Db 279 -----PNMTSKADAL-DWMLLYTDSKRTNIGVYLNOKPLDPTTMSAPTLLAT 328
Qy 367 GPCGYSDLAAL-----ENEGFLGCFEFCGKQCEQIAFLRTDREILSHVOG 414
Db 329 GICAYSDLQNHGHPGCGSPQSGCLYE-----SNNYEIVELMTLQAFPAVFG 377

RESULT 5
T34999
probable neuramidase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T34999
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1998
A:Reference number: 221564
A:Accession: T34959
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-648 <SEE>
A:Cross-references: EMBL:AL034443; PIDN:CAA22361.1; GSPDB:GN00070; SCOEDB:SC4B5.07c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC4B5.07c

Query Match 11.0% Score 259; DB 2; Length 648;
Best Local Similarity 25.3% Pred. No. 3e-14;
Matches 128; Conservative 61; Mismatches 143; Indels 174; Gaps 30;
Qy 5 TSCSFS-----SPLFOQE-----DKRG-VYRIPALYVPPAHTFLAFA 42
Db 15 TSCALAVSPFPAHARPGDRAPAGEQVLFDAARDPGGYACFRIPAIVRTTGDG-TLLAFA 73
Qy 43 EKR-SSSKDEDAHLVLRGLRTGOSVOMEPLKSLMKATLPGHRTMPCPWERKSGVY 101
Db 74 EGRVLDCAADGGIDIVLRRSLDGGRT--WGPLRVVNDG--GGDTHGNPAPVVDRTGRVL 129
Qy 102 LFFICVGHVTERQOIMSGRNPRLCFT-C-----SODAGYSWSDVRLTEVIGPE 152
Db 130 LL-----ETYNAGRTSDACVPCARVPHVQHSDDGGRTWSAPRDLSPILPPD 178
Qy 153 VTHWATFVAGPGHGIQL-----OSGRLIIPAYA-----YYIPFWFCFRLPYRARP- 198

Db 179 WNSW--YATGPVHGQVLTGGAHPGRLVGVVNAETWDCERSEMGPV-----PA 223
Qy 199 -----HSLMIYSDDLGATWHHGRIL-----IKPMVTVECEVAEYIGRAGH 237
Db 224 GGRVRTANHAALVVDGGEHWTGATDTPWAADGTFROKPSLTLTERAD----- 276
Qy 238 PVLVCSARTPN-----RIHRAEALSDHGECEFOKPVLSHQLECEPHGCGSVVFCPLIPG 293
Db 277 GALLVSGREENGTDPGHRTQALSDGGDSFAAPFA-----LP- 314
Qy 294 GCQDLAGEDAPAIQOSPLLCSSVRPEAGTSLSEWLLYSHPTNKKRRVDLGIYLNOSPL 353
Db 315 ---DLY---APQVQGAVLRLG-----NRLLSAPADPDRRRMTV--RSSRD 353
Qy 354 EAACW-----SRPWILHCGPGYSDLAALNEGLFGCLFECGKQCEQIAF-RLF 403
Db 354 GGATWDSADRGTVVTRDW-----AGYSDLVTVDDTV-GLLYEGGKTARDEIRFARLT 406
Qy 404 TDREILSHVOG-DCSTPCM--NSEPS 426
Db 407 ADR--LAPPRGPDPTTDLAANAAPA 430

RESULT 6
A45244
exo-alpha-sialidase (EC 3.2.1.18) - Micromonospora viridifaciens
N:Alternate names: neuraminidase
C:Species: Micromonospora viridifaciens
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 22-Oct-1999
C:Accession: A45244
R:Sakurada, K.; Ohta, T.; Hasegawa, M.
J. Bacteriol. 174, 6896-6903, 1992
A:Title: Cloning, expression, and characterization of the Micromonospora viridifaciens
A:Reference number: A45244; MUID:93015752
A:Accession: A45244
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-647 <SAK>
A:Cross-references: GB:D01045; NID:g216782; PIDN:BAA00852.1; PID:d1001316; PID:g21678
A:Note: sequence extracted from NCBI backbone (NCBIP:116820)
C:Keywords: glycosidase; hydrolase

Query Match 8.6% Score 201.5; DB 2; Length 647;
Best Local Similarity 24.0% Pred. No. 2.7e-09;
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;
Qy 12 PLFOQED-----KRGV-TYRIPALYVPPAHTFLAFABKRSKDEDAHLVLR-RRGLRT 64
Db 49 PLYTEQDLAVNGREGFPNRYIPALT-VTPDGLLASVDGRPTGIDAPGPNILQRRSTDG 107
Qy 65 GQSVQWEPLKSLMKATLPGHRTMPCPWERKSGVYVYLFFICVGHVTERQOIMSGRNP 124
Db 108 GRTWGEQVVSAGQTTAPIKGFSDFSLVDRETGTIFNF-----HVSYQKGAGSRPG 161
Qy 125 -----RLCFCISODAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQLO----- 171
Db 162 TDPADPNVLHANVATSDGGLTWSH-RTITAD-ITPD-PGHSRFAASGEQIQLRYGPHA 218
Qy 172 GRL-----IIPAYAYIPIFWFCFRLPYRAPHLSMIYSDDLGATWHHGRILI-----KPMV 222
Db 219 GRLIQOVTIINAAGAF-----QAVSVYSDDHGRTRWAGEAVGVGMDENK 262
Qy 223 TVECEVAEYIGKAGHPVLVYCSARTPNHRAEALSDHGECEFOKPVLSHQLECEPHGCGOS 282
Db 263 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGGHSGYGPVTIDRDLDPPTN--NAS 313
Qy 283 VVSFCPLIPGCGODLAGEDAPAIQOSPLLCSSVRPEAGTSLSEWLLYSHPTNKKRRV 342
Db 314 IIRAFP-----DAP-----AGSARAKVLLFNSAASQTSR- 342
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGPGCYSDLAALNEGLFGCLFECGKQCEQ 397

Db	343	-----SQGTIRMSCDDGQTWPVSKVFPQGSMSYSTITALP-DGTYGLLYERGT-----G	390
Qy	398	IAPFLFTDREILSHVQDC---STPGMNSPSKK	428
Db	391	IRVANFN-----LAWLGGICAPFTIPDVALEPGQQ	420

RESULT 7
T37097
probable secreted neuraminidase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37097
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21588
A:Accession: T37097
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-748 <SNU>
A:Cross-references: EMBL:AL109950; PIDN:CAB52948.1; GSPDB:GN00070; SCOEDB:SCJ4.14C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCJ4.14C

Query Match	8.4%	Score	197	DB 2	Length	748			
Best Local Similarity	24.9%	Pred. No.	7.9e-09						
Matches	103	Conservative	56	Mismatches	153	Indels	102	Gaps	21
Qy	22	VYRIPALIIYVPPAHTTFLAFAEKR-SSSKQEDALHLVLRRLRTGQSVQVOWEPLKSLMKAT	80						
Db	76	VCFRAPAVVKAADG-TLLAFAGEGRIGSCDPAFIDIVVKRYVNGA----	W\$ALQVVAH\$S	130					
Qy	81	LPGHRTMPCVPWERKSGYVLFYTCVQGHVTERQOIMSGRNPAPLRCIFCSODAGY\$WSD	140						
Db	131	-AGHIYHNVTVPVDAASGRVVLYTENYDHTHR-----	IASEDDGLHWTA	174					
Qy	141	VRDLTEEVIGVETHWAT-----	FAVGPCHQIQL----	QSGRLIIPAVYIPIEFW	187				
Db	175	ADDSADV-----	W\$TAWGALLVAGQMATPASAQIUTGHRHAGRLVAG-----	217					
Qy	188	FCFRLPYRAPHSL---	MIYDDLGATWHHGRLL---	IKPMYTVCEVAEVIKAGHPVLY	241				
Db	218	MTVRVAPCAAPANLGGALIIYSDGGLTWRLCASSLGAEPVAGAQ-ELS-	LFRGDDGSLF	274					
Qy	242	CSAR-----	T\$NRHRA-EALSIDHGECEQKP-VLSHQICEPHPCQGSV\$VFCLEIPGG	294					
Db	275	VTARNEEGSDTRAVYAVSGDQGL\$TSDFALLPMDLPGTGQASTLALREKNR\$G-	333						
Qy	295	QODLAGEDAPAIQQ\$PLICS\$VRPEPEAGT\$E\$WLLY\$HPTNKKRRRVLDGLIYNQ\$PLE	354						
Db	334	-----	YDRALEAPVGNREDLITR\$F-----	D\$GLTW-QDRAAD	367				
Qy	355	AAC\$SRPWILHCGPCGYSDLAALENEGLFCGLFCG\$TQKECEQIA\$RFL\$T\$DREI	408						
Db	368	GA-----	LVRDQ\$Y\$Y\$S\$MTVL--GGG\$FGLIY\$EAGT\$K\$Y\$ODIRFA\$FT\$EADL	414					

```

RESULT      8
NMCLSS
exo-alpha-sialidase (EC 3.2.1.18) precursor - Clostridium septicum
N:Alternate names: neuraminidase
C:Species: Clostridium septicum
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C:Accession: S15994
R:Rothe, B.; Rothe, B.; Roggentin, P.; Schauer, R.
Mol. Gen. Genet. 226, 190-197, 1991
A:Title: The sialidase gene from Clostridium septicum: cloning, sequencing, expression in
A:Reference number: S15994; MUID:91238693
A:Accession: S15994
A:Molecule type: DNA
A:Residues: 1-1014 <MOL>

```

A: Cross-references: EMBL:X63266; NID:q40662; PIDN:CAA4916.1; PID:q40663
C: Superfamily: Clostridium exo-alpha-sialidase
C: Superfamily: Clostridium exo-alpha-sialidase
F: 1-26/Domain: signal sequence #status predicted <SIG>
F: 27-1014/Product: exo-alpha-sialidase #status predicted <WAT>

Query Match	7.98;	Score 186;	DB 1;	Length 1014;
Best Local Similarity	27.5%;	Pred. No. le-07;		
Matches	83;	Conservative	37;	Mismatches 102; Indels 80; Gaps 16;
Qy	126	LCFTCSODAGYSWSDVRDLTEEVIGPEVTHWATF-AVGPGHGIIQIQ-----SGRLIIPAYA	180	
Db	560	LSLIYSSDDGGTWDSPIDLNKEV-----KTDWRFLGTGPGKCHQIKTCRYAGRLLPVYL	615	
Qy	181	YYIPFECFRLPYRAPHSLMIYSSDDLGAATW-----HGRLI-----KPMVT---	223	
Db	616	TNAGSF-----QSSAVIYSSDDNGATWNIGETATDGLMDNGDRASAEITTTNTS	664	
Qy	224	-----VECEVAEIVIGRAGHPVLYCSARTPNRHAE-ALSIDHGECEKPKVLSHQICEPP	276	
Db	665	GGVGQLTECGQVEM--PNGQLKMFMRNTGGNSGRVRIATSDGGATWEDDDVVVRDENIKEP	722	
Qy	277	HGCGSVVSCFPLETPGSCQDLAGEDAPAIQGSPLICSSVRPEAGTLSSEWLLYSHPT	336	
Db	723	Y-CQLSVINY-----SQIDGRDA-----IIFAIPDAN-----YPNRV	754	
Qy	337	NKKRRVDL-----GIYLNOSPLEAACWSRPWILHCGPCGYSLAALENEGLCGLIFE-CGT	391	
Db	755	NGTVRVGLITENGSENGEPRYDIEWRYNKVVAPCTYGYSCLSMPN-GEIGLFEVGRGS	813	
Qy	392	KQ	393	
Db	814	RO	815	

RESULT 9
 140866
 exo-alpha-sialidase (EC 3.2.1.18) - Clostridium perfringens
 C:Species: Clostridium perfringens
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
 C:Accession: I40866
 R:Traving, C.; Schauer, R.; Roggentin, P.
 Glycoconj. J. 11, 141-151, 1994
 A:Title: Gene structure of the 'large' sialidase isoenzyme from Clostridium perfringens
 A:Reference number: I40865; MUID:95102306
 A:Accession: I40866
 A:Status: preliminary; translated from GE/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-694 <RES>
 A:Cross-references: EMBL:X87369; NID:g853805; PIDN:CAA60796.1; PID:g853807
 C:Genetics:
 A:Gene: nanH
 C:Keywords: glycosidase; hydrolase

	Query Match	7.3%;	Score 172;	DB 2;	Length 694;	
	Best Local Similarity	20.8%;	Pred. No.le=06;			
	Matches 102; Conservative	65;	Mismatches 143;	Indels 180;	Gaps 25;	
Qy	24 YRIPALIVVPAHTFLAFAEKRSSKD--EDALHLVLR--GLRT---GOSVOVEPLK- 74		: : : : : : : : :			
Db	265 YRIPALFTKGTGLIASIDARRHGADAPNNDITAVRRSEGGKTWEGQLIMDYPDKS 324		: : : : : : : : :			
Qy	75 SLMKATLPGHTRMPCPYWERKSGYVILF-----FCVC- 107	: :	: : : : : : : : :			
Db	325 SVIDTTL-----IQDDETGRIFLLVTHFPSPKYGFWNAGLCSGFKNIDGKEYLCLY 374	: :	: : : : : : : : :			
Qy	108 -----QGHYTER-----QQIMSGRNPAR-----LCFI 129	: :	: : : : : : : : :			
Db	375 DSSGKEFTVRENVVVDKGNKTEVTYNALGDLFKNNGTKIDINNSTAPLKAGTSYINLV 434	: :	: : : : : : : : :			
Ov	130 CSQDAGYSWSVDRLDEVIPEVTHEWTATF-AVGPGHGIIQLQS-----GRLIIPAYAYVIP 184	: :	: : : : : : : : :			

```

Db 435 YSDDDGRTWSEPONINFQV-----KKDMMKFLGIAPGRGQIQRNGEHKGRIVPVVY----- 485
Qy 185 FWEFCFRLPY---RAPHSLMIYSDDLGATW-----HHGRLLIK-----PM 221
Db 486 -----YTNEKGROSSAVIYSDSGKNWTIGESPNDNRKLENGKIINSKTLSDDAQP 536
Qy 222 VTVECEVAEIVGAGHPVLVYCSARTPNRIRAEALSTDHGCECFQKPVLSHQICEPHGCGQ 281
Db 537 LT-ECQVVEN--PNGQKLP--MRNLSGYLNIATSPDGGATWDETVEKDTNVLEPY-CQL 590
Qy 282 SVVSFCEPLPIPGGCQDLADGAPAIQOSPILCSSLVSRPEAGTSLSESWLLYSHPTNKKR- 340
Db 591 SVINL-----SQKVDGKDA-----VFISNPNARSRS 616
Qy 341 ----RVDL----GYLNSQSPLEAACWSRPWILHCGPCGYSDLAALENEGLFCFLFECGTK 392
Db 617 NGTVRIGLINQVGYENGEPKYEFDKNKLVKPGYYAVSCLTSLN-GNIGLLYECPIS 675
Qy 393 QECEQIAFRL 402
Db 676 EEMSYLEMNL 685

RESULT 10
A49227
sialidase - Actinomyces viscosus
C:Species: Actinomyces viscosus
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 22-Oct-1999
C:Accession: A49227
R:Yeung, M. K.
Infect. Immun. 61, 109-116, 1993
A:Title: Complete nucleotide sequence of the Actinomyces viscosus T14V sialidase
A:Reference number: A49227; MUID:93114861
A:Contents: T14V
A:Accession: A49227
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-901 <YEU>
A:Cross-references: GB:L06898; NID:Q289087; PIDN:AAA21932.1; PID:gl141852
A>Note: sequence extracted from NCBI backbone (NCBIN:121598, NCBIPI:121599)

```

```

Query Match          7.2%; Score 169; DB 2; Length 901;
Best Local Similarity 23.0%; Pred. No. 2.5e-06;
Matches          91; Conservative 50; Mismatches 155; Indels 100; Gaps

Qy 24 YRPALIVVPANTFLFAFK-----RSSKKDEDAHLHLVLRGLTGQSVQWEPKLSMK 178
      ||||: |||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 310 YRPAITAPNGDLLISYDERPDKNNGSGDAPNPNHIVQRRTSDGGKT--WSAPTYIHQ 367
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 79 ATLPGHRT--MNPQPVWERKSGVYLFFCIVOGHVTEROQ-----IMSGRNPARLCFI--- 129
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 368 GTEGKKVGSYDSYVVDHOTGTFINP-----HVKSYYQGHGSGRGDTPNRRGIIQAE 421
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 130 --CSODAGYSWSVDRLTEEVIGPETH--WATFVAGPGHGIQIQ-----SGRLIIPAYA 180
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 422 VSTSDNGWTW-----THRTITADITDKKPWTARFAASQGGIQQHGHGPHACRLV----- 470
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 181 YYTPFWFFCFRLPYARP-----HSLMIYSDDLGLATWHHGRLIKPMVTVECEVAEVIKKA 235
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 471 -----QQYTIRTAGGVAQVSVYSDDHGKTMQAG--TPGTGMDENKVKVELSD 516
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 236 GHPVLYCSARTPNRRHRAEALSIDHGECFOKPYLSHOLCEPHPHCCQGSVVSFCPLPIGCC 295
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 517 GSLMLNSRSDGSGFRKVAHSTDGGTGWSEPVSKDNLPSVDONAQ--IIRAFNPAAAP----- 571
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 296 QDLAGEDAPAIQOSPLICSSVRPEP---BAGTLESWLLYSHPNTNKRVRDLGIYLNQSP 352
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 572 -----DDPRAKVL-LLSHSPNRPNSRDRGTIS-----MSC 601
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Qy 353 LEACWSRPNILHCGPGYSDLAALNEGLFQCLFE 388
      ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:

```

Db 602 DDGASWTSKVFHEPFGVYTTI-AVQSDGSIGLLSE 636

RESULT 11

S20590

exo-alpha-sialidase (EC 3.2.1.18) - *Actinomyces viscosus*

C:Species: *Actinomyces viscosus*

C:Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 22-Oct-1999

C:Accession: S20590

R:Henningsen, M.; Roggentin, P.; Schauer, R.
Biol. Chem. Hoppe-Seyler 372, 1065-1072, 1991

A:Title: Cloning, sequencing and expression of the sialidase gene from *Acti*

A:Reference number: S20590; MUID:92162190

A:Accession: S20590

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-913 <HEI>

A:Cross-references: EMBL:X62276; NID:g39254; PIDN:CAA44166.1; PID:g39255

C:Keywords: glycosidase; hydrolase

Query Match 6.6%; Score 154; DB 2; Length 913;
Best Local Similarity 23.5%; Pred. No. 5.1e-05;
Matches 94; Conservative 48; Mismatches 150; Indels 108; Gaps 20;

[illegible]

RESULT 12
T30287
exo-alpha-sialidase (EC 3.2.1.18) - Streptococcus pneumoniae
N:Alternate names: neuraminidase
C:Species: Streptococcus pneumoniae
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T30287
R:Camara, M.; Boulnois, G.J.; Andrew, P.W.; Mitchell, T.J.
Infect. Immun. 62, 3688-3695, 1994
A:Title: A neuraminidase from Streptococcus pneumoniae has the features of
A:Reference number: Z20807; MUID:94341870
A:Accession: T30287
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1035 <CAM>
A:Cross-references: EMBL:X72967; NID:g587550; PID:g587553; PIDN:CAAS1473.1
C:Genetics:
A:Gene: nana
C:Keywords: glycosidase; hydrolase

	Query Match	6.3%; Score 149; DB 2; Length 1035;
	Best Local Similarity	19.6%; Pred. No. 0.00016;
	Matches 103; Conservative	63; Mismatches 159; Indels 200; Gaps 22;
QY	9 FSSPLFQEDRGV-TYPALIVPPAHTFLAFNAEKSSKDEDAHLVLRR---- <td> : : : :</td>	: : : :
Dd	330 FESGRNGPNKDIGKSYIPALLTKGTLTAGADERLHSDGDIGMVRSEDNGKT 389	: : : : : : : : :
QY	64 TQSQVQEPLKSLMKATLPG-----HRTMNPCPVWE----- 94	: : : : : : : :
Dd	390 WDRVTITNLRDNPKASDPSGPSVINDMVLVDQPETKRIFSIYDMFFPEGKIFGMSSQK 449	: : : : : : : :
QY	95 -----RKSGYVYLFFICVQGHVTERQOISGRNPA---- 124	: : : : : : : :
Dd	450 EEAYKKIDGKTYILYREGKGAVTIRENGTVY---TPDGKATDYRVVDPVKPAYSDK 505	: : : : : : : :
QY	125 -----RLCFTCSQDAGYSWSDVRDLTEEVIGPEVTH 155	: : : : : : : :
Dd	506 GDLYKGNOLLGNIFYTTNKTSFPFRIAKDSYLWMSYSDDDGKTSAPQDITPW----KAD 561	: : : : : : : :
QY	156 WATP-AVGPGHGIIQLS---GLRIIPAY----AYIPFWFEFCRLPYRAPPHLSMIYSD 206	: : : : : : : :
Dd	562 WKFLGVGGPGTGIVLRNGPHGRILIPVYTNNVSHL-----NGSQSRRIIYSD 610	: : : : : : : :
QY	207 DLGATWHHRGLIKPMVTVECEVAEVIKAGHPVLYCSARTNHRAE-----ALSIDHGE 261	: : : : : : : :
Dd	611 DHGKTHAGEAVN-----DNROVDGOKIH-----SSTNNRRQAQNTSTVVQLNNGD 657	: : : : : : : :
QY	262 CFQRPVLVSHLCPEPHCGQGVSVFCPLPIPGGCODLAGEDA-----PAIQOSP 310	: : : : : : : :
Dd	658 -----VKLFMRGLTGDLQVATSKDGGVTWEKDKIRYPOVKDVY 695	: : : : : : : :
QY	311 LLCSSVRPEAGTLESWLLYSHTNKKR-----RVDLG---IYLNOSPLEAACWS 359	: : : : : : : :
Dd	596 VOMSAITHME-----GREYIILSNAGGPKRNGMVHVARVEENGELTWLKHNPLOK 747	: : : : : : : :
QY	360 RPWLHLCGPCGYDLALENEGFGCLFECGTK-QECEQIAFLRF 403	: : : : : : : :
Dd	748 -----GEFAYNSLQELGN-GEYGILYEHTKEGQNAYLTLSFRF 784	: : : : : : : :
RESULT	13	
G98063	exo-alpha-sialidase (EC 3.2.1.18) precursor (neuraminidase A) [imported] - Streptococcus pneumoniae	
C:Species:	Streptococcus pneumoniae	
C:Date:	22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001	
C:Accession:	G98063	
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McArthur, Y. P.; Sun, P.-M.; Winkler, M.E.		
J. Bacteriol.	183, 5709-5717, 2001	
A:Authors:	Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskun, A.; Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.	
A:Reference number:	A97872; PMID:21429245; PMID:11544234	
A:Accession:	G98063	
A>Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-1035 <CUR>	
A:Cross-references:	GB:AEO07317; PIDN:AAL00340.1; PID:g15459200; GSFPDB:GN00174	
C:Genetics:		
C:Keywords:	glycosidase; hydrolase	

Query Match 6.3%; Score 149; DB 2; Length 1035;
Best Local Similarity 19.6%; Pred. No. 0.00016;
Matches 103; Conservative 63; Mismatches 159; Indels 200; Gaps 22;

QY 9 FSPFLQEQDKRGV-TYRIPALIVYVPPAHTFLFAEKRSSKDEDLHLVLR----GLR 63
Db 330 FESGRNGKPNKDGKFSYRIPALLTKDGTGLIAGADERLRLHSSDWDGDTGVIIRSEDNKGT 389

[illegible]

A; Molecule type: DNA
A; Residues: 1-545 <AKI>
A; Cross-references: DDBJ:D28493
A; Note: In this paper, the nucleotides GAC were missed between DNA sequence 1220 and s' translated sequence, Ala instead of Gly- Pro
R; Russo, T.A.; Thompson, J.S.; Godoy, V.G.; Malamy, M.H.
J Bacteriol. 172, 2594-2600, 1990
A; title: Cloning and expression of the Bacteroides fragilis TAL2480 neuraminidase gene
A; Reference number: A35264; MUID:90236922
A; Accession: A35264
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 239-250; 318-328, 'A'; 379-390; 426-437; 486-497 <RUS>
A; Cross-references: GB:M31663
C; Genetics:
A; Gene: nan H
C; Keywords: glycosidase; hydrolase
F; 1-2/Domain: signal sequence #status predicted <SIG>
F; 23-545/Product: exo-alpha-stalase #status predicted <MAT>
F; 241-248, 320-327, 381-388, 428-435, 488-495/Region: Asp box
F; 241-248, 320-327, 381-388, 428-435, 488-495/Region: Asp box

Query Match 6.2%; Score 146; DB 2; Length 545;
Best Local Similarity 21.5%; Pred. No. 0.00013;
Matches 84; Conservative 59; Mismatches 157; Indels 90; Gaps 16;
ov 24 YRIPALITYVPPAHTFLAFACKRSSKD-EDALHLVLRRLRGRTGOSVOWE----PLKSLMK 78

Db 203 FRIPGLVTNKGTLGVDVRYNSVDLQEHVDVGLSRSTDGKT--WEKMRPLPAFGEF 260
Qy 79 ATLPG--HRTMNPVWERKSGYVLFICVQGHVTER-----QQIMSGRNPALRCFICS 131
Db 261 GGLPAGONGVGDPSILVDTKTNVYVAAWTHGMGNORAMWSSHPGMDMHTAQLVLAKS 320
Qy 132 ODAGYSMSDVR-DLTEEYIGPEVTHWATFAVPGHGIQLOQSGRLIIPAYAYVYIPFWFFCF 190
Db 321 TDDCKTWSGPPINITEQVKDPS--WYFLQGPGRGITSMDGLVFPQF----- 367
Qy 191 RLPYRAPHSLMIYSDDLGATWHHGRLIKPMVTVECEVAEVIGKAGHPVLVCSARTPNR- 249
Db 368 -IDSTRVPNAGIMYSKDGKKNKKNHYARTN-TTEAQVAEV-----EPGVLMNLNMRNRG 420
Qy 250 -HRAEALSIDHGECFOKPVLSHOLCEPHGCGSVSFPCLEIPGCGODLAGEDAPAIQ 308
Db 421 GSRVAITKOLGKTWTEHESRK-----ALPE 447
Qy 309 SPLICSSVRPEPEAGTLESWLLYSHPTNKKRRVLDLGIYLNQSPLEAACWS----- 359
Db 448 SVCNASLISYKAKDNLGKDLLEFSPNPTTKGRYNTIKISLD--GGVTNSPEHQLLIDE 505
Qy 360 -RPWILHCGPCGYSDLAALENEGLFGCLFE 388
Db 506 GNNW-----GYSCLSMIDKETI-GILYE 527

RESULT 15
A37234
exo-alpha-sialidase (EC 3.2.1.18) - Clostridium sordellii
C:Species: Clostridium sordellii
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 17-Mar-2000
C:Accession: A37234
R:Rothe, B.; Roggentin, P.; Frank, R.; Bloecker, H.; Schauer, R.
J. Gen. Microbiol. 135, 3087-3096, 1989
A:Title: Cloning, sequencing and expression of a sialidase gene from Clostridium sordellii
A:Reference number: A37234; MUID:90132537
A:Accession: A37234
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-404 <ROT>
A:Cross-references: GB:M31584; NID:g144910; PIDN:AAA23280.1; PID:g144911
C:Superfamily: trypanastigote-specific surface antigen
C:Keywords: glycosidase; hydrolase

Query Match 6.2%; Score 145.5; DB 2; Length 404;
Best Local Similarity 22.1%; Pred. No. 0.0001;
Matches 80; Conservative 47; Mismatches 122; Indels 113; Gaps 16;

Qy 24 YRIPALIVPPAHTFLAPAEKR-SSSKDEADALHLVRLRGLTGQSVQWEPKLSMKATLP 82
Db 54 FRIPSLQTLADG-TMLAFSDIRYGAEDHAYIDICAASKSTGNDGTQDWTYKVMENDRIDST 112
Qy 83 GHRTMNPV-----WERKSGYVYLFVICVQGHVTERQRIQIMSGRNPALCFI 129
Db 113 FSRVNDSTTVYDTRGRIILIAGSNKKNWNA-----SSTSLRSDWSVOMV 158
Qy 130 CSQDAGYSWSDVRDLT--EEVIGPEVTHWATFAVPGHGIQLOQSGRLIP-----A 178
Db 159 YSDNGETWSKVDLTNNKARIKQPSNTIGLAGVSGIVMSDGTIVMPIQIALRENN 218
Qy 179 YAYIPIFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLIKPMVTVECEVAEVIGKAGHP 238
Db 219 NNY-----SSVIYKDNGETWTMGNKVPDPKTSNNVIELDG----- 256
Qy 239 VLYCSARTPNR-HRAEALSIDHGECFOKPVLSHOLCEPH-----CGQSVVSF-- 286
Db 257 ALIMSSRNDGKNRYASYISYDMGSTW-----EYDPLHNKISTGNSGCGQSGFIKVTA 309
Qy 287 -----CPLEIPGG-----CQDLAGEDAPAIQOQSLICSSVRPEAGT 324

Db 310 KDGHRGLGFISAPKNWTKGGYVRDNTVYVIMIDFDDL-----SKGIRE-----LCS-----PYPEDGN 359
Qy 325 LS 326
Db 360 SS 361

Search completed: October 7, 2002, 14:59:51
Job time: 73 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:59:58 ; Search time 17.45 seconds
(without alignments)
949.683 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSPFLQEDRR.....LSHVQDCSTPGMNSPSSK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	NER3_BOVIN	Q7859 bos taurus
2	1916.5	81.6	428	NER3_HUMAN	Q9uq49 homo sapien
3	1518	64.7	418	NER3_RAT	Q99pw5 rattus norv
4	1501	63.9	418	NER3_MOUSE	Q9jmh7 mus musculu
5	711.5	30.3	380	NER2_HUMAN	Q9y3r4 homo sapien
6	705	30.0	379	NER2_MOUSE	Q9jmh3 mus musculu
7	705	30.0	379	NER2_RAT	Q84627 rattus norv
8	692	29.5	379	NER2_CRIGR	Q64393 cricetus
9	255.5	10.9	415	NER1_HUMAN	Q99519 homo sapien
10	248.5	10.6	409	NER1_RAT	Q99pw3 rattus norv
11	245.5	10.5	409	NER1_MOUSE	Q35657 mus musculu
12	201.5	8.6	647	NANH_MICVI	Q02834 micromosp
13	186	7.9	1014	NANH_CLOSE	P29767 clostridium
14	149	6.3	1035	NANH_STRPN	Q59959 streptococ
15	145.5	6.2	404	NANH_CLOSO	P15698 clostridium
16	128.5	5.5	266	NANH_BACFR	P31206 bacteroides
17	128	5.5	382	NANH_CLOPE	P10481 clostridium
18	107	4.6	1877	PCK5_MOUSE	Q04592 mus musculu
19	105	4.5	781	NANH_VIBCH	P37060 vibrio chol
20	99.5	4.2	2768	THYG_HUMAN	P01266 homo sapien
21	98	4.2	931	NRP2_MOUSE	Q35375 mus musculu
22	97	4.1	925	NRP2_RAT	Q35276 rattus norv
23	94.5	4.0	1216	AREGP_RAT	Q63191 rattus norv
24	94	4.0	381	NANH_SALTY	P29768 salmonella
25	94	4.0	2282	ZAN_RABIT	P57999 oryctolagus
26	91	3.9	697	NANH_STRPN	Q34727 streptococ
27	89.5	3.8	1255	PER2_HUMAN	Q15055 homo sapien
28	89	3.8	493	AMR_DRDEL	Q9njp0 drosophila
29	89	3.8	525	UL32_EBV	P03184 Epstein-bar
30	87	3.7	493	1 AMR_DROME	Q84808 drosophila
31	87	3.7	493	1 AMR_DROTK	Q77018 drosophila
32	86.5	3.7	931	NRP2_HUMAN	Q60462 homo sapien
33	85	3.6	794	FURI_HUMAN	P09958 homo sapien

34	85	3.6	918	1 KPCM_MOUSE	Q62101 mus musculu
35	85	3.6	1025	1 CR2_MOUSE	P19070 mus musculu
36	84.5	3.6	731	1 SUFL_XENLA	Q9pun2 xenopus lae
37	84.5	3.6	1187	1 PTNE_HUMAN	Q15678 homo sapien
38	84.5	3.6	1255	1 LEM2_HUMAN	P04626 homo sapien
39	84	3.6	484	1 LEM2_PIG	P98110 sus scrofa
40	84	3.6	493	1 AMYR_DROAN	P18344 drosophila
41	84	3.6	2437	1 NOTC_BRARE	P46530 brachydanio
42	83.5	3.6	1141	1 SRE2_HUMAN	Q12772 homo sapien
43	83.5	3.6	1433	1 Y310_HUMAN	Q15027 homo sapien
44	83	3.5	493	1 AMYR_DROYA	Q76264 drosophila
45	83	3.5	494	1 AMYR_DROAP	Q77011 drosophila

ALIGNMENTS

RESULT 1					
NER3_BOVIN					
ID NER3_BOVIN	STANDARD;	PRT;	428 AA.		
AC	Q97859;				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)				
DE	(N-acetyl-alpha-neuraminidase 3).				
CN	NEU3.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.				
RC	TISSUE=Brain;				
RX	MEDLINE=99143165; PubMed=9988745;				
RA	Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,				
RA	Sawada M.;				
RT	"Molecular cloning and characterization of a plasma membrane-				
RT	associated sialidase specific for gangliosides.";				
RL	J. Biol. Chem. 274:5004-5011(1999).				
CC	- FUNCTION: Plays a role in modulating the ganglioside content of				
CC	the lipid bilayer at the level of membrane-bound sialyl				
CC	glycoconjugates.				
CC	- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,				
CC	alpha-(2->8)-glycosidic linkages of terminal sialic residues in				
CC	oligosaccharides, glycoproteins, glycolipids, colominic acid and				
CC	synthetic substrates.				
CC	- SUBCELLULAR LOCATION: Membrane-associated (By similarity).				
CC	- TISSUE SPECIFICITY: Expressed in brain.				
CC	- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.				
CC	- SIMILARITY: CONTAINS 3 BNR REPEATS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: AB08184; BAA75071.1; -				
DR	InterPro: IPR002860; BNR.				
DR	Pfam: PF02012; BNR; 3...				
KW	Hydrolase; Glycosidase; Membrane; Repeat.				
REPEAT	129 140				
FT	REPEAT 129 140				
FT	REPEAT 203 214				
FT	REPEAT 254 265				
FT	REPEAT 254 265				
FT	SITE 24 27				
FT	ACT_SITE 25 25				
FT	ACT_SITE 45 45				
FT	ACT_SITE 50 50				
FT	ACT_SITE 87 87				
FT	POTENTIAL.				
FT	POTENTIAL.				

```
FT ACT_SITE 225 225 POTENTIAL.
FT ACT_SITE 245 245 POTENTIAL.
FT ACT_SITE 341 341 BY SIMILARITY.
FT ACT_SITE 371 371 POTENTIAL.
FT ACT_SITE 388 388 POTENTIAL.
SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F21 CRC64;

Query Match 100.0%; Score 2348; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.5e-201;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVTSFSSPLFQEDKRGVYRIPALIVPPAHTFLFAEKRSSSKDEDALHLVLR 60
DB 1 MEEVTSFSSPLFQEDKRGVYRIPALIVPPAHTFLFAEKRSSSKDEDALHLVLR 60
QY 61 GLRTGQSVQVEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFVFCVQGHVTERQQIMSG 120
DB 61 GLRTGQSVQVEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFVFCVQGHVTERQQIMSG 120
QY 121 RNPALRCFCISQDAGYSWSDVRDLTEVIGPETHWATFAVPGHGIGLOSGRLIIPAYA 180
DB 121 RNPALRCFCISQDAGYSWSDVRDLTEVIGPETHWATFAVPGHGIGLOSGRLIIPAYA 180
QY 181 YYIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEYIGKAGHPVL 240
DB 181 YYIPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEYIGKAGHPVL 240
QY 241 YCSARTPNRHRAREALSDHGECKQKPVLSHOLCEPHGCGSVVFCFLEIPGCGQDLAG 300
DB 241 YCSARTPNRHRAREALSDHGECKQKPVLSHOLCEPHGCGSVVFCFLEIPGCGQDLAG 300
QY 301 EDAPAIQOSPILCSSLVPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNOSPLEAACWSR 360
DB 301 EDAPAIQOSPILCSSLVPEPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNOSPLEAACWSR 360
QY 361 PWILHCGPCGYSDLAALENGLFCGCLPECTKQCEQIAFRFTDRILSHVQDCSTPG 420
DB 361 PWILHCGPCGYSDLAALENGLFCGCLPECTKQCEQIAFRFTDRILSHVQDCSTPG 420
QY 421 MNSEPSKK 428
DB 421 MNSEPSKK 428

RESULT 2
NER3_HUMAN
ID NER3_HUMAN . STANDARD; PRT; 428 AA.
AC Q9UQ49; Q9NOE1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99335353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
RT sialidase.";
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RA Croci G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
```


Db	301	QDAPTQQSS	-PGSSURLNEEEAGTSPESWLLYSHPTSRKQRVDLGIYLNQTPLEAACWSR	359
Qy	361	PWLHCGPCGYSDLAALNEGLFGCLFECGTQKECQIAFRFLTQREILSHVQGDCTPG	420	
Db	360	PWLHCGPCGYSDLAALNEGLFGCLFECGTQKECQIAFRFLTQREILSHVQGDCTSPG	419	
Qy	421	MNSEPSK	427	
Db	420	RN--PSQ	424	
RESULT 3				
ID	NER3_RAT	NER3_RAT		
AC	Q99PW5	STANDARD;	PRT: 418 AA.	
DT	01-MAR-2002	(Rel. 41, Created)		
DT	01-MAR-2002	(Rel. 41, Last sequence update)		
DT	01-MAR-2002	(Rel. 41, Last annotation update)		
DE	Slialidase 3 (EC 3.2.3.18) (Membrane slialidase)	(Ganglioside slialidase)		
DE	(N-acetyl-alpha-neuraminidase 3).			
GN	NEU3			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
ON	NCBI_TaxID=10116;			
OX	11			
PC	SEQUENCE FROM N.A.			
RP	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RX	MEDLINE=21092669; PubMed=11162381;			
RT	Hasegawa T., Feijoo Carnero C., Wada T., Itoyama Y., Miyagi T.;			
RT	"Differential expression of three slialidase genes in rat			
RL	development.";			
RL	Biochem. Biophys. Res. Commun. 280:726-732(2001).			
CC	-1- FUNCTION: Plays a role in modulating the ganglioside content of			
CC	the lipid bilayer at the level of membrane-bound sialyl			
CC	glycoconjugates (By similarity).			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-			
CC	alpha-(2->8)-glycosidic linkages of terminal sialic residues in			
CC	oligosaccharides, glycoproteins, glycolipids, colominic acid and			
CC	synthetic substrates.			
CC	-1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).			
CC	-1- TISSUE SPECIFICITY: Expressed in brain, cardiac muscle and weakly			
CC	in liver.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.			
CC	-1- SIMILARITY: CONTAINS 3 BNR REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstat-			
CC	ion at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	way modified and this statement is not removed. Usage by and for commer-			
CC	c entities requires a license agreement (See http://www.isb-sib.ch/anno-			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AB026841; BAB32440.1; -.			
DR	InterPro: IPR002860; BNR.			
DR	Pfam: PF02012; BNR: 3			
KW	Hydrolase; Glycosidase; Membrane; Repeat.			
FT	REPEAT 129 140			
FT	REPEAT 201 212			
FT	REPEAT 252 263			
FT	SITE 24 27			
FT	ACT_SITE 25 25			
FT	ACT_SITE 45 45			
FT	ACT_SITE 50 50			
FT	ACT_SITE 87 87			
FT	ACT_SITE 223 223			
FT	ACT_SITE 243 243			
FT	ACT_SITE 339 339			
FT	ACT_SITE 369 369			
FT	ACT_SITE 386 386			
SQ	SEQUENCE 418 AA; 46980 MW; 7CC46F2E5952E240 CRC64;			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AB026842; BAA92868.1; -
DR MGI: 1355305; Neu3.
DR InterPro: IPR002860; BNR.
DR Pfam: PF02012; BNR: 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 FRIP MOTIF.
FT ACT_SITE 25 25 BY SIMILARITY.
FT ACT_SITE 45 45 POTENTIAL.
FT ACT_SITE 50 50 POTENTIAL.
FT ACT_SITE 87 87 POTENTIAL.
FT ACT_SITE 223 223 POTENTIAL.
FT ACT_SITE 243 243 POTENTIAL.
FT ACT_SITE 339 339 BY SIMILARITY.
FT ACT_SITE 369 369 POTENTIAL.
FT ACT_SITE 386 386 POTENTIAL.
SQ SEQUENCE 418 AA; 46846 MW; 64853FC963FE7686 CRC64;

Query Match 63.9%; Score 1501; DB 1; Length 418;
Best Local Similarity 67.3%; Pred. No. 8.6e-126;
Matches 282; Conservative 47; Mismatches 86; Indels 4; Gaps 2;

QY 1 MEEVTSFSSPLFOQEDKRGVYRIPALIVYPPAHTFLAFAEKRSKDEDAHLVLR 60
DB 1 MEEVPPYLSSTLFFQEQSGVYRIPALIVYPPAHTFLAFAEKRTSVRDEDAACLVLR 60
QY 61 GLRTGQSVQWEPKLSLMKATLPGHRTMPCPVWERKSGVYVFFICVGHVTERQOIMSG 120
DB 61 GLMKRSVQWGPORLLMEATLPGHRTMPCPVWEKNTGRVYVFFICVGHVTERCQIVWG 120
QY 121 RNPALCFCTCSODAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQSLGRLIIPAYA 180
DB 121 KNAARLCFLCSEDAGCSWGEVKDLTEEVIGSEVKNWATFVAGPGHGTQLHSGRLIIPAYA 180
QY 181 YIIPFFWFCFLPYRAPHSLMIYSDDLGATWHHGRILKPMVTVECEVAEVIKAGHPVL 240
DB 181 YVYSRWFCLFAC--SVKPHSLMIYSDDFGVTHHGKFIQVPTGECQVAEAGTAGNPVL 238
QY 241 YCSARTPNRHRAREALSDHGECEQKPKVLSHOLCEPPHGCQSVVFCPLPFGCCQDLAG 300
DB 239 YCSARTSPRFAEAFSTDSGCCFQKTLNQLHPRTGCGSVVSRFLKMPNTYQDSIG 298
QY 301 EDAPAIQSPLLCSSVRPEAGTLESWLLYSHPTNKKRRVDGLYLNOSPLEAACWSR 360
DB 299 KGAPATQKCPLLDSPLEVEKGAETSPATWLLYSHPTSKRRINLGIYNNRPLEVNCWSR 358
QY 361 PWILHCGPCGYSLAALENGLFCGECCTKQECQIAFRLETDRILLSHVQGDSTP 419
DB 359 PWILNRGPGSYDLAAVEODLVACLFECGERNEYERIDFLCFLSDEVLVS--CEDCTSP 415

RESULT 5
NER2_HUMAN STANDARD; PRT; 380 AA.
AC Q9Y3R4;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-neuraminidase 2).
GN NEU2.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=99208670; PubMed=10191093;
RA Monti E., Preti A., Rossi E., Ballabio A., Borsani G.;
RT "Cloning and characterization of NEU2, a human gene homologous to
RT rodent soluble sialidases.";
RL Genomics 57:137-143(1999).
CC -!- FUNCTION: Hydrolyzes sialylated compounds.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
CC alpha-(2-8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, fetal liver and
CC embryonic carcinoma cell line NT2B1.
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Y16535; CAB41449.1; -
DR MIM: 605528; -
DR InterPro: IPR002860; BNR.
DR Pfam: PF02012; BNR: 2.
KW Hydrolase; Glycosidase; Repeat.
FT REPEAT 127 138 BNR 1.
FT REPEAT 197 208 BNR 2.
FT SITE 20 23 FRIP MOTIF.
FT ACT_SITE 21 21 BY SIMILARITY.
FT ACT_SITE 46 46 POTENTIAL.
FT ACT_SITE 243 243 POTENTIAL.
FT ACT_SITE 304 304 BY SIMILARITY.
FT ACT_SITE 334 334 POTENTIAL.
FT ACT_SITE 355 355 POTENTIAL.
SQ SEQUENCE 380 AA; 42230 MW; 9D18F1041A2D4F44 CRC64;

Query Match 30.3%; Score 711.5; DB 1; Length 380;
Best Local Similarity 40.7%; Pred. No. 9.6e-56;
Matches 170; Conservative 59; Mismatches 114; Indels 75; Gaps 16;

QY 10 SSPLFOQED--KRGV--TYRIPALIVYPPAHTFLAFAEKRSKDEDAHLVLR 63
DB 3 SLPLVQKESVFGCAHAYRIPALLYLPQOSLLAFEAQRAKDEHAELIVLRGVDAP 62
QY 64 TGSQVQWEPKLSLMKATLPGHRTMPCPVWERKSGVYVFFICVGHVTERQOIMSGRN 123
DB 63 THQ-VQWQAEVVAQARLDGHRSMNCPPLYDAQTGTLFFIAIPGVTEQQQLQTRNV 121
QY 124 ARLCFTCSODAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQL--QSGRLIIPAY 181
DB 122 TRLCQVTSIDHGTWSSPRDLTDAALGPAYREWSSTFVAGPGHGLCLRSRLVWPAY 181
QY 182 Y-----IPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRILKPMVTVECEVAEVI 233
DB 182 RKLHPTQRIIPS-AFCF-----LSHDGRTWARGHFV-AQDTLEQCAVEV-- 224
QY 234 KAGHP-VLYCSARTPNRHRAREALSDHGECEQKPKVLSHOLCE--PPHGCQSVVFCPL 291
DB 225 ETGEQRVVTILNARSHLRARVQAQSTNDLDFQESQLVKLVPEPPQCGQSVISF----- 279
QY 292 PGCCQDLAGEDAPAIQOSPLLCSSVRPEAGTLESWLLYSHPTNKKRRVDGLYLNQ 350
DB 280 -----PSPRSGPGSPAQWLLYTHPTHSWRADLGLAYNP 313

FT	ACT_SITE	21	21	BY SIMILARITY.
FT	ACT_SITE	46	46	POTENTIAL.
FT	ACT_SITE	243	243	POTENTIAL.
FT	ACT_SITE	303	303	BY SIMILARITY.
FT	ACT_SITE	333	333	POTENTIAL.
FT	ACT_SITE	354	354	POTENTIAL.
FT	CONFLICT	40	40	K -> R (IN REF. 1).
FT	CONFLICT	91	97	YKQTKT -> WTSKKD (IN REF. 1).
SEQ	SEQUENCE	379 AA:	42403 MW:	02124A46398F6793 CRC64;
Query Match 30.0%; Score 705; DB 1: Length 379;				
Best Local Similarity 39.8%; Pred. No. 3.6e-55;				
Matches 164; Conservative 63; Mismatches 117; Indels 68; Gaps				
QY	12	PLFQOED--KRGV-TYRIPALIIYVPAHTLAFAPAEKRSSKSDALHLVLRRL--RTGO	66	
DB	5	PVLQKETLFTGTGHAVRIPALLYLKKQKTLLLAFAPAEKRASKTDSHAELIIVLRGSGYNEATN	64	
QY	67	SVQWEPLKSLMKATLCGHRTPNCPWVERKSGVYVLFICVQGHVTERQOIMSGRNPRL	124	
DB	65	RVKMQPEVVYTAQLEGRSMNCPDYDKTKTLFLFFTAVGRVSEHQLHTKVNWRL	124	
QY	127	CFICSQDAGYSWSDVRDLTEEVIGPEVTHWATFVAGPGHGIIQIQ--SGRLIIPAYAY--	182	
DB	125	CCVSSDTHGRTWSPIDLTETTIGSTHQAETAVFGHCLQRLNPAGSLIVPAYAYRKL	184	
QY	183	-----IPFWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVTGKAG	236	
DB	185	HPAQKPTPF-AFCF-----ISLDHGHTWKLGNFVAEN-SLEQVAEV-GTGA	238	
QY	237	HPVLYCSARTPNRHRAEALSIDHGECFOKPVLSHOLCEPPHGGCGSVVFCPLIEPGCGQ	296	
DB	229	QRWVYLNAEFLGARVQAQSPNDGLDFQDNRVVSCLKVLEPPHGGCGGSVAE-----	278	
QY	297	DLAGEDAPATQQSPLGSSVRRPEAGTUSESWLLYSHPTNKKRRVLDLGYTLNQSPLEAA	356	
DB	279	-----HNPI-----SKPHA---LDTWLLYTHPTDSNRNLTGLVYLNQMLDPT	318	
QY	357	CWSRPWLLHCGPCGYSLAAL-----ENEGFLGCLFCGCTKQCEQIATFLFT	404	
DB	319	AWSEPTLLAMGICAYSDLQNMGGPGSGPFGCLYESG---NYEEIIFLIPT	367	
RESULT	7			
NER2_RAT	NER2_RAT	STANDARD;	PRF;	379 AA.
AC	Q64627; Q63705;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-			
DE	neuraminidase 2).			
GN	NEU2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;			

RT cytosolic sialidase gene.";
RL Glycobiology 5:511-516(1995).
CC -!- FUNCTION: Hydrolyzes sialylated compounds.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D16300; BAA03805.1; -;
DR EMBL; D50606; BAA09169.1; -;
DR InterPro: IPR002860; BNR.
DR Pfam; PF02012; BNR; 2.
KW Hydrolase; Glycosidase; Repeat.
FT REPEAT 127 138 BNR 1.
FT SITE 197 208 BNR 2.
FT SITE 20 23 FRIP MOTIF.
FT ACT_SITE 21 21 BY SIMILARITY.
FT ACT_SITE 46 46 POTENTIAL.
FT ACT_SITE 243 243 POTENTIAL.
FT ACT_SITE 303 303 BY SIMILARITY.
FT ACT_SITE 333 333 POTENTIAL.
FT ACT_SITE 354 354 POTENTIAL.
SQ SEQUENCE 379 AA; 42382 MW; 55583C7043CA9784 CRC64;

Query Match 30.0%; Score 705; DB 1; Length 379;
Best Local Similarity 39.5%; Pred. No. 3.6e-55;
Matches 167; Conservative 60; Mismatches 126; Indels 70; Gaps 13;
QY 12 PLFQOE---DKRGVTRIPALIVVPPAHTFLAFAPKRSKDEDAHLVLRGLRTGQS- 67
DB 5 PVLQKETLPHTEVYAYRIPALLYLKKQKTLAPFAEKRSKRTDEHAEILVLRGSGYNGATN 64
QY 68 -VOWEPLKSLMKATLPGHRTMPCPVWKRSGVYLFFFCVQGHVTERQOIMSGRNPRL 126
DB 65 HVKMQPEEVVTOQLGHSRMNCPPLYDKQTKTLFLFFIAPGVRVSEQHQLQTRVNVTRL 124
QY 127 CFTCSODAGYSWSDRLTEEVIGPEVTHWATFVAGPGHGIQL--OSGRLLIIPAYAY-- 182
DB 125 CRVSTDYGMNNSPVQDLTETTTIGSTHQDWATFVAGPGHCLQLRNRAGSLVPAYAYRKL 184
QY 183 -----IPWFECFLPYRAPHSLMIYSDLLGATWHHGRLLKPMVTVECEVAEVIKAG 236
DB 185 HPVHKPTFP-ACFC-----ISLDHGHWTGELGNFVSEN-SLEQCAVEV-GTGA 228
QY 237 HPVLYCSARTPNRRAEALSIDHGEQFQPVLSHOLCEPHGCGQSWSF-CPLEIPGCG 295
DB 229 HRVVLNARSFTGARVQAQSPNDGLDFQDNQVVKLVEPPHCHGSGVAFHFTSKPDCL 288
QY 296 QDLGADAPAIQOOSPLLCSSVRPEPAGTSLSESWLYSHPTNKKRVDLGIYLNQSPLEA 355
DB 289 RHVAA-----YTHPTDSRNTNLGVYLNQTPLDP 317
QY 356 ACSWRPWILHCGCGYSDLA--ALENEG--LFGCLFECTKQCEQIAFLRTDREILSH 411
DB 318 TAWSEPTLATGTCAYSIDLQIWGLDGPSPQFGCLYESGNYDE---IIFLMTLQAAFT 374
QY 412 VQG 414
DB 375 VHG 377

RESULT 8
NER2_CRIGR
ID ID NER2_CRIGR STANDARD; PRT; 379 AA.
AC Q64393;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sialidase 2 (EC 3.2.1.18) (Cytosolic sialidase) (N-acetyl-alpha-
DE neuraminidase 2).
GN NEU2.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC CCL61;
RX MEDLINE=95036975; PubMed=7949662;
RA Ferrari J., Harris R., Warner T.G.;
RT "Cloning and expression of a soluble sialidase from Chinese hamster
RT ovary cells: sequence alignment similarities to bacterial
RT sialidases.";
RL Glycobiology 4:367-373(1994).
CC -!- FUNCTION: Hydrolyzes sialylated compounds.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -!- SIMILARITY: CONTAINS 2 BNR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U06143; AAA19746.1; -;
DR InterPro: IPR002860; BNR.
DR Pfam; PF02012; BNR; 2.
KW Hydrolase; Glycosidase; Repeat.
FT REPEAT 127 138 BNR 1.
FT REPEAT 197 208 BNR 2.
FT SITE 20 23 FRIP MOTIF.
FT ACT_SITE 21 21 BY SIMILARITY.
FT ACT_SITE 46 46 POTENTIAL.
FT ACT_SITE 243 243 POTENTIAL.
FT ACT_SITE 303 303 BY SIMILARITY.
FT ACT_SITE 333 333 POTENTIAL.
FT ACT_SITE 354 354 POTENTIAL.
SQ SEQUENCE 379 AA; 41962 MW; B5AFFBC6B6BE88B1 CRC64;

Query Match 29.5%; Score 692; DB 1; Length 379;
Best Local Similarity 38.6%; Pred. No. 5.1e-54;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;
QY 13 LFQQEDKRGVTRIPALIVVPPAHTFLAFAPKRSKDEDAHLVLRGLRTGQS--VQW 70
DB 12 LFQTDG---YAYRIPALIVLSKQKTLAPFAEKRLTKTDEHADLFVLRGSGYNADTHQVQW 68
QY 71 EPLKSLMKATLPGHRTMPCPVWKRSGVYLFFFCVQGHVTERQOIMSGRNPRLCFLC 130
DB 69 QAEVYVTOAYLEGRHSMSPCLYDKQTKTLFLFFIAPGVRVSEQHQLQTRVNVTRLCHIT 128
QY 131 SODAGYSWSDRLTEEVIGPEVTHWATFVAGPGHGIQLQ--SGRLIIPAYAYIIPFWFF 188
DB 129 STDHKTWSAVQDLTDTTIGSTHQDWATFVAGPGHCLQLRNTAGSLVPAYAYR----- 182

ID	NANA_STRPN	STANDARD;	PRT;	1035 AA.
AC	Q59599;	Q54722;		
DT	15-DEC-1998	(Rel. 37, Created)		
DT	15-DEC-1998	(Rel. 37, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Sialidase A precursor	(EC 3.2.1.18) (Neuraminidase A).		
DE	NANA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R36A / NCTC 10319;			
RX	MEDLINE=94311870; PubMed=8063384;			
RA	Camara M., Boulnois G.J., Andrew P.W., Mitchell T.J.;			
RA	"A neuraminidase from Streptococcus pneumoniae has the features of a			
RT	surface protein.";			
RL	Infect. Immun. 62:3688-3695(1994).			
RL	[2]			
RN	SEQUENCE OF 882-1035 FROM N.A.			
RP	STRAIN=SEROTYPE 6;			
RX	MEDLINE=96326329; PubMed=8759848;			
RA	Berry A.M., Lock R.A., Paton J.C.;			
RA	"Cloning and characterization of nanB, a second Streptococcus			
RT	pneumoniae neuraminidase gene, and purification of the NanB enzyme			
RT	from recombinant Escherichia coli.";			
RT	J. Bacteriol. 178:4854-4860(1996).			
CC	-1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL			
CC	NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.			
CC	-1- SIMILARITY: CONTAINS 4 BNR REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed, usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X72967; CAA51473.1; -			
DR	EMBL; U43526; AAC44391.1; -			
DR	HSP; Q02834; 1EUR.			
DR	InterPro; IPR002860; BNR.			
DR	InterPro; IPR001791; Laminin_G.			
DR	InterPro; IPR004124; sialidase_N.			
DR	Pfam; PF02012; BNR; 4.			
DR	Pfam; PF02973; sialidase_N; 1.			
DR	SMART; SM00282; LamG; 1.			
KW	Hydrolase; Glycosidase; Cell wall; Repeat; Signal.			
FT	SIGNAL 1 ? POTENTIAL.			
FT	CHAIN 1 ? 1035 SIALIDASE A.			
FT	SEQUENCE 1035 AA; 114741 MW; C5B8A2D7A12E12F3 CRC64;			
SC	-----			
Query Match	6.38; Score 149; DB 1; Length 1035;			
Best Local Similarity	19.68; Pred. No. 2.7e-05;			
Matches 103; Conservative	63; Mismatches 159; Indels 200; Gaps 22;			
QY	9 FSSPLFQOEKRGV-TYRIPALIVVPPAHTFLAFAEKRSSKEDALHLVLR----	GLR 63		
DB	330 FESGRCKPNKDGKISYRIPALLTKDGTGLIAGADERLHSSDWDGDIWMIRSEDNGKT 389			
QY	64 TQGSVQWEPLKSLMKATLPG-----HRTMFPVWE-----	94		
DB	390 WGDRTVTTNLRDNPKASDPSIGSPVNDIMVLVDPETKRIFSYIDMPEKGIFGMSSQK 449			
QY	95 -----RKSGVYVLFCTVOGHVTEROQIISGRNPA----	124		
DB	450 EAYKKIDGKTYILYREGEKGYATIRENGTV----TPDGKATDYRVVVDVVKRAYS DK 505			

Qy	125	-----R	LCFCISQDAGYSWSDVRDLTEEVIGPEVTH	151
Db	506	GDLYKGNOLLGNIYFTTNKTSPPFI	AKDSYLWMSYSDDDGKTWSPQDITPMV	561
Qy	156	WATF-AVGP	HGHIQLOS---GRLLIIPAY---AYIIPWF	206
Db	562	WKFLGVGPGGIVLRNGPHKGRILLI	PYVNTNNVSHL---NGSQSRIIYSD	610
Qy	207	DLGATWHHGRLIKPMWTVCEVAE	VIGKAGHPVLYCSARTPNRHRAE	261
Db	611	DHGKTHAGEAVN-----DNRQ	VGQKIH-----SSTMNRR	657
Qy	262	CFQKPVLSHQLCEPHGCGQSV	SFVCFPLEIPGCGQDLAGEDA	310
Db	658	-----VKLF	MGRGLTGDLQVATSKDGGTV	695
Qy	311	LLCSVRPEPAGT	LSBSWLLYSHPTNKKR-----RVDLG	359
Db	696	VOMSAIHTMHE-----GREY	ILTSNAGGPKRENGMVHILARVEENGE	747
Qy	360	RPTWLHCPCGYSDLAAL	ENEGFLGCFEGCTK-QECEQIA	403
Db	748	-----GEFAYNSLQELGN	-GEYGILYEHTKQCNAYTLS	784
RESULT 15				
NANH_CLOSO				
ID	NANH_CLOSO	STANDARD;	PRT;	404 AA.
AC	PI5698;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Sialidase precursor (EC 3.2.1.18) (Neuraminidase).			
OS	Clostridium sordellii.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1505;			
RP	[1]			
RN	SEQUENCE FROM N. A., AND SEQUENCE OF 28-45.			
RC	STRAIN-G12;			
RC	MEDLINE=90132537; PubMed=2693593;			
RA	Redhe B., Roggentin P., Frank R., Bloecker H., Schauer R.;			
RT	"Cloning, sequencing and expression of a sialidase gene from			
RT	Clostridium sordellii G12.";			
RL	J. Gen. Microbiol. 135:3087-3096(1989).			
CC	!- FUNCTION: SIALIDASES HAVE BEEN SUGGESTED TO BE PATHOGENIC FACTORS			
CC	IN MICROBIAL INFECTIONS.			
CC	!- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL			
CC	NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.			
CC	!- SUBCELLULAR LOCATION: Periplasmic.			
CC	!- PTM: IT IS POSSIBLE THAT THE SIALIDASE IS CLEAVED IN FRONT OF A			
CC	CYSTEINE WITHIN THE LEADER PEPTIDE, FORMING A GLYCERIDE			
CC	THIOETHER BOND WHICH LINKS THE PROTEIN TO THE MEMBRANE. A SECOND			
CC	PROTEOLYTIC CLEAVAGE RELEASES THE MATURE EXTRACELLULAR PROTEIN.			
CC	!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.			
CC	!- SIMILARITY: CONTAINS 5 BNR REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstat-			
CC	ion at the European Bioinformatics Institute. There are no restrictions on			
CC	the use by non-profit institutions as long as its content is in no			
CC	modified and this statement is not removed. Usage by and for commer-			
CC	cial entities requires a license agreement (See http://www.isb-sib.ch/annou			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M31584; AAA23280.1; -			
DR	PIR; A37234; A37234.			
DR	HSSP; P29768; 2SIL.			
DR	InterPro; IPR002860; BNR.			
DR	Pfam; PF02012; BNR; 5.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.			
DR	Hydrolase; Glycosidase; Periplasmic; Signal; Repeat.			
KW	SIGNAL			
FT	1			
FT	27			

```

Query Match      6.3%; Score 149; DB 1; Length 1035;
Best Local Similarity 19.6%; Pred. No. 2.7e-05;
Matches 103; Conservative 63; Mismatches 159; Indels 200; Gaps 22;

QY      9  FSSPLFQOEKRGV-TYRIPALIVYPAHFTLPAERKRSSKDEDLHLVLRR-----GLR 63
      | | : : : : : | : | : | : | : | : | : | : | : | : | : | : |
Db      330  FESGRNGKPNKGDKSYRIPALLTKDGTGLIAGADERLHSSDWDGIMVIRREDNGKT 389
      | | : : : : : | : | : | : | : | : | : | : | : | : | : | : |

QY      64  TQGSVQNEPLKSLMKAILPG-----HRTWNPFCPWE----- 94
      | | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      390  WGDRTVTITLNRDNPKASDPGSPVNDIMVLVDOPETKRIFSYDMFPEFGKIGFMSSQK 449
      | | : : : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      95  -----RKSGVYVLFVCVOGHVYTEROOIMSGRNPA----- 124
      | | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      450  EAYYKIDGKTYQILYREGKGYATIRENGFTV-----TPDGKATDYRVVVVDPKPAYSDK 505

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:59:18 ; Search time 50.13 Seconds

(without alignments)
1476.996 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTSFSFSLFQEDKR.....LSHVQGDCTPGMNSPSKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp_plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp_unclassified.*
- 15: sp-virus.*
- 16: sp-bacteriap.*
- 17: sp-archaeop.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	11.0	648	2 Q9ZBW1	Q9Zbw1 streptomyc
2	197	8.4	748	2 Q9SIV9	Q9Siv9 streptomyc
3	172	7.3	694	2 Q59310	Q59310 clostridium
4	169	7.2	901	2 Q44562	Q44562 actinomyc
5	154.5	6.6	544	2 Q45145	Q45145 bacteroides
6	154	6.6	913	2 Q59164	Q59164 actinomyc
7	150.5	6.4	404	2 P94675	P94675 clostridium
8	142.5	6.1	404	2 P94674	P94674 clostridium
9	140.5	6.0	404	2 P94676	P94676 clostridium
10	137.5	5.9	1003	2 Q9AHP2	Q9Ahp2 arcanobacte
11	132.5	5.6	762	5 Q27701	Q27701 macrobact
12	125	5.3	740	16 Q97Q99	Q97q99 streptococ
13	123	5.2	568	12 Q91II6	Q91i16 frog adenov
14	120	5.1	382	2 Q59311	Q59311 clostridium
15	117.5	5.0	773	2 P77848	P77848 clostridium
16	115.5	4.9	643	5 Q9BHJ5	Q9Bhj5 trypanosoma

17	113.5	4.8	642	5 Q26969	Q26969 trypanosoma
18	113.5	4.8	771	5 Q9GSF0	Q9gsf0 trypanosoma
19	111.5	4.7	642	5 Q26968	Q26968 trypanosoma
20	109.5	4.7	642	5 Q26967	Q26967 trypanosoma
21	109.5	4.7	771	5 Q9GU83	Q9gu83 trypanosoma
22	109.5	4.7	964	5 Q26963	Q26963 trypanosoma
23	109.5	4.7	1070	2 Q9EZV7	Q9ezv7 pasteurella
24	108.5	4.6	642	5 Q26966	Q26966 trypanosoma
25	108	4.6	1170	2 Q9AJR8	Q9ajr8 erysipelothe
26	107.5	4.6	517	12 Q9YUS1	Q9yus1 turkey aden
27	106.5	4.5	879	5 Q00773	Q00773 trypanosoma
28	106.5	4.5	1060	5 Q26964	Q26964 trypanosoma
29	103	4.4	556	4 Q9UMP5	Q9ump5 homo sapien
30	102	4.3	680	11 Q55001	Q55001 mus musculu
31	101.5	4.3	700	5 Q08672	Q08672 trypanosoma
32	100	4.3	567	4 Q9BWE0	Q9bwe0 homo sapien
33	100	4.3	567	4 Q9BUZ6	Q9buz6 homo sapien
34	95.5	4.1	389	10 Q23780	Q23780 chlamydomon
35	95.5	4.1	558	12 Q84169	Q84169 oliveros vi
36	94.5	4.0	2820	5 Q9VLT6	Q9vit6 drosophila
37	92	3.9	567	4 Q9NZH2	Q9nzh2 homo sapien
38	92	3.9	628	4 Q9BTF0	Q9btf0 homo sapien
39	92	3.9	841	11 P97484	P97484 mus musculu
40	92	3.9	1080	16 Q9CM43	Q9cm43 pasteurella
41	91.5	3.9	700	5 Q26965	Q26965 trypanosoma
42	91	3.9	1299	12 Q9YUJ1	Q9ytj1 ateline her
43	90	3.8	381	12 Q91GK2	Q91gk2 epiphyas po
44	90	3.8	1436	10 Q9LFY6	Q9lfy6 arabidopsis
45	90	3.8	1572	5 Q44938	Q44938 haemochus

ALIGNMENTS

RESULT 1

Q9ZBW1 ID Q9ZBW1 PRELIMINARY; PRT; 648 AA.

AC Q9ZBW1; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE PUTATIVE NEURAMIDASE.

GN SC4B5.07C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Seeger K.J., Harris D.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kieser H.M., Denapante D., Eichner J., Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

DR EMBL: AL034443; CAA22361.1; -.

DR HSSP: Q02834; 1EUR.

DR InterPro: IPR002860; BNR.

DR InterPro: IPR001791; Laminin_G.

DR Pfam: PF02012; BNR: 5.

DR SMART: SM00282; LamG: 1.

SQ SEQUENCE 648 AA; 68216 MW; F1CD5835276679EB CRC64;

```

Query Match      11.0%; Score 259; DB 2; Length 648;
Best Local Similarity 25.3%; Pred. No. 1.4e-16;
Matches 128; Conservative 61; Mismatches 143; Indels 174; Gaps 30;

QY 5 TSCSFS-----SPLFOE-----DKRG-VYTRIPALIVPPAHTFLAFA 42
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 15 TSCALAVSPFFAHARPGRAPAGEQVLEDAARDPGGYACFRIPAIYRTTGD-TLLAFA 73
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 43 EKR-SSSKDEDALHLVLRRLTGOSQVOWEPLKSLMKATLPGHRTMPCPVWERKSGVYV 101
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 74 EGRVLDCAADGDIDIVLRSLDGGRT--WGPLRVVNDG--GGDTHGNPAPVVDRTATGRVL 129
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 102 LFFICVQGHVTRQOIMSGRNPRLCFI-C-----SQDAGYSWSVDRLITEEVIGPE 152
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 130 LL-----ETYNAGRTSDSACVPCARVPHVQHSDDGGRTWSAPRDLSPILPPD 178
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 153 VTHWATFVAGPGHGIQL-----QSGRLIIPAYA-----YIPFWFCFRLPYRARP- 198
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 179 WNSW--YATGPVHGVLGGHAGPGLVGVNAETWDGERSEMGVP-----PA 223
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 199 -----HSLMIYSDDLGATWHHGR-----KPMVTVECEVAEIVGKAGH 237
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 224 GGWRVTANHAALVYSDDGGHWRGTATWPVAADGTFROKPSLTLTERAD----- 276
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 238 PVLVCSARTPN-----RHRAEALSDHGEFCQKPVLSHOLCEPPHGCQSGSVVFCPLTPG 293
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 277 GALLVSGREENGTDPGHRTQALSRDGGDSFAAPFA-----LP- 314
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 294 GQDLAGEDAPAIQSPILCSSVRPEPEAGTLESWLLYSHPTNKKRRVLDGLIYNOSPL 353
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 315 ---DLV---APVOGAVLRLG-----NRILLSAPADPRRTMTV--RSSRD 353
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 354 EAACW-----SRPWILHCGPGCYSDLAALNEGLGCLFECGTQKECEQIAF-RLF 403
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 354 GGATWDSADRGTVYTRDW-----AGISDLYTVDDDTV-GLLYEGGRTDARDEIRFARLT 406
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 404 TDREILSHVQG-DCSTPGM--NSPES 426
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 407 ADR--LAPPRGPDPTPDLAANAAPA 430
   ||| : : : : : | : : : : | : : : : | : : : : |

RESULT 2
Q9S1V9 PRELIMINARY; PRT; 748 AA.
AC Q9S1V9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE SECRETED NEURAMINIDASE.
GN SCJ4-14C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C.; Harris D.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D.; Parkhill J.; Barrell B.G.; Rajandream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M.; Kleser H.M.; Denapaita D.; Eichner A.; Cullum J.;
RA Kinashi H.; Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL; AL109950; CAB52948.1; -.
DR HSSP; Q02834; 1EUT.
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR; 4.
SQ SEQUENCE 748 AA; 79060 MW; 20E829A5AB867571 CRC64;

Query Match      8.4%; Score 197; DB 2; Length 748;
Best Local Similarity 24.9%; Pred. No. 1.8e-10;
Matches 103; Conservative 56; Mismatches 153; Indels 102; Gaps 21;

QY 22 VTYRIPALIYPPAHTFLAFAEKR-SSSKDEDALHLVLRRLTGOSQVOWEPLKSLMKAT 80
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 76 VCFRAPAVVKAADG-TLIAFAEGRIGSCDSTASIDIVVKRYVNGA----WSALQVVARHS 130
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 81 LPGAHTMPCPVWERKSGVYVLFVICVQGHVTRQOIMSGRNPRLCFICQDAGYSWS 140
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 131 -AGHIYHNVTVPVDAASGRVVLYTENYDHIR-----IASEDDGLHWA 174
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 141 VRDLTEEVIGVETHWAT-----FVGPFGHGIQL-----QSGRLIIPAYIIPFWF 187
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 175 ADDISADV-----WSTAWGALYAGQMATGPASAIQLTHGRHAGRLVAG----- 217
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 188 FCFRLPYRAPHSL---MIYSDDLGATWHHGR-----KPMVTVECEVAEIVGKAGHPVLY 241
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 218 MTVRVAPCAAPANLGGALIYSDGGLTWRLGASSILGAEPVGAQ-ELS--LPERGDSILF 274
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 242 CSAR-----TPNRRA-EALSDHGEFCQKPVLSHOLCEPPHGCQSGSVVFCPLIEPG 294
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 275 VTARNEEGSDTRAVYAVSGDGLSFTSDFALLPMDLFGTIGQASTLALREKNRDG- 333
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 295 QCDLAGEDAPAIQSPILCSSVRPEPEAGTLESWLLYSHPTNKKRRVLDGLIYNOSPLE 354
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 334 -----YDRALFAAPVGNREDLTIRESF-----DGLTW-QDAAD 367
   ||| : : : : : | : : : : | : : : : | : : : : |
QY 355 AACRSRPWILHCGPGCYSDLAALNEGLGCLFECGTQKECEQIAFRFTDREI 408
   ||| : : : : : | : : : : | : : : : | : : : : |
Db 368 GA-----LVKDGYSAYSSMTVL-GGDTFGILYEAGTKQYQDIRFATFTTEADL 414
   ||| : : : : : | : : : : | : : : : | : : : : |

RESULT 3
Q9S310 PRELIMINARY; PRT; 694 AA.
AC Q9S310;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EXO-ALPHA-SIALIDASE PRECURSOR (EC 3.2.1.18).
GN NANH.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=A99;
RA MEDLINE=95102306; PubMed=7804004;
RA Traving C.; Schauer R.; Roggentin P.;
RT "Gene structure of the 'large' sialidase isoenzyme from Clostridium
RT perfringens A99 and its relationship with other clostridial nanH
RT proteins.";
RL Glycoconj. J. 11:141-151(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=A99;
RA Schauer R.;
RT "Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RL EMBL; X87369; CAA60796.1; -.
DR HSSP; Q02834; 1EUT.
DR InterPro; IPR002860; BNR.
DR InterPro; IPR004124; sialidase_N.
```

DR Pfam: PF02012; BNR: 5.
SQ SEQUENCE 901 AA; 92861 MW; BCF6369DB6DF7E57 CRC64;

Query Match 7.2%; Score 169; DB 2; Length 901;
Best Local Similarity 23.0%; Pred. No. 1.2e-07;
Matches 91; Conservative 50; Mismatches 155; Indels 100; Gaps 18;

QY 24 YRIPALTYVPPAHTFLAFAEK-----RSSSKDEADLHLVRLRRGLRTGOSVOWEPLKSLMK 78
| | | | :
Db 310 YRIPAATTAPNGDLLISYDERPKDNGSGDAPNPNHIVORSTGGKT--WSAPTYIHQ 367
| | | | :
QY 79 ATLPQHRT--MNPCEPWERKSGYVLFFTCVOGHVTEROQ-----TMSGRNPARLCFI--- 129
| | | | :
Db 368 GTETGRKVGSDPSYVDVHQGTGFENF-----HVKSVDQGSGRGDTDPENRGIIQAE 421
| | | | :
QY 130 --CSODAGYSWSDVRDLTEEVIGPEVTH---WATFAVPGIHCIOLO---SGLIIPAYA 180
| | | | :
Db 422 VSTSTONGTW-----THRTITADITDKPWTAFAASGGGIOIGHPHAGRLV----- 470
| | | | :
QY 181 YYIPFWFCFRPYRAPR-----HSLMIYSDDLGATWHHGRLIKPMVTVECEVAEVIKA 235
| | | | :
Db 471 -----QQYTIRTAGGAOVAVSVDHDKTQWAG---TPIGTCMDENKVVELSD 516
| | | | :
QY 236 GHVPVLYCSARTNNRHRAEALSIDHGECKFKPVLSHOLCEPHGCOSVVSVFCLEIPGCC 295
| | | | :
Db 517 GSLMLNSRASDGSGFRKVAHSTDGGQTWPSEVPVSDKNLPDSVNAQ--IIRAPPNAAP--- 571
| | | | :
QY 296 QDLAGEDAPAIOQSPILCSSVRPEP---EAGTLSESLLSYSHPTNKRRVDLGIVLNOSP 352
| | | | :
Db 572 -----DDPRAKYL-LLSHPNRPMSRDRTIS-----MSC 601
| | | | :
QY 353 LEACSRPWLHCGPCGYSDLAALNEGLFGCLPE 388
| | | | :
Db 602 DQGSMTTSKFHEFPVGVTTI-AVSDGSIGLLSE 636
| | | | :

RESULT 5
Q45145 PRELIMINARY; PRT; 544 AA.

ID Q45145 PRELIMINARY; PRT; 544 AA.
AC Q45145;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEURAMINIDASE PRECURSOR.
GN NANH.
OS Bacteroides fragilis.
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
OC Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YCH46;
RX MEDLINE=94380076; PubMed=8093075;
RA Kimoto S., Ono T., Tsutsui H., Kinouchi T., Kataoka K., Ohnishi Y.;
RT "Complete sequence of the Bacteroides fragilis YCH46 neuraminidase-
RT encoding gene";
RL Biochem. Biophys. Res. Commun. 203:914-921(1994).
RN [2]
RP SEQUENCE OF 1-78 FROM N.A.
RC STRAIN=TM4000;
RA Godoy V.G., Dallas M.M., Gallegos R.A., Malamy M.H.;
RT "Neuraminidase Gene Expression in Bacteroides fragilis TM4000 Is
RT Transcriptionally Regulated and Induced by N-acetyl Neuraminic Acid.";
RL Submitted (Oct-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; D28493; BAA05853.1; -;
DR EMBL; AF031639; AAD01937.1; -;
DR HSSP; P29768; 3SIL.
DR InterPro; IPR002860; BNR.
DR Pfam; PF02012; BNR; 5.
KW Signal.
FT CHAIN 1 22 POTENTIAL.
FT SIGNAL 23 544 NEURAMINIDASE

DR Pfam: PF02012; BNR: 4.
DR Pfam: PF02973; sialidase.N; 1.
KW Signal; Hydrolase; Glycosidase.
FT SIGNAL 1 41 POTENTIAL.
SQ SEQUENCE 694 AA; 77327 MW; 22D23E25A0A8DC72 CRC64;

Query Match 7.3%; Score 172; DB 2; Length 694;
Best Local Similarity 20.8%; Pred. No. 4.5e-08;
Matches 102; Conservative 65; Mismatches 143; Indels 180; Gaps 25;

QY 24 YRIPALTYVPPAHTFLAFAEKRSSKD--EDALHLVLR--GLRT---GQSVOWEPLK- 74
| | | | :
Db 265 YRIPALFKTEGLIASIDARRHGGADAPNDIDTAVRSEDGGTWDGEGTIIMDYDPSK 324
| | | | :
QY 75 SLMAKATLPGHRTNCPVWERKSGYVLF-----FICV- 107
| | | | :
Db 325 SVIDTTL-----IQDETGRIFLLVTHFPSKYGFVNAGLSGFGKNIDCKEYLCY 374
| | | | :
QY 108 -----QGHVTER-----QQIMSGRNPAP-----LCFI 129
| | | | :
Db 375 DSSGKEPTVENVYDKDNKTEYTNALCDLFGKTKIDNINSYAPLAKAGTSYNLV 434
| | | | :
QY 130 CSODAGYSWSDVRDLTEEVIGPEVTHWF-avgphgqlols----grLIIPAYAYIP 184
| | | | :
Db 435 YSDDDGKTWSEPQNINQV---KKDMKELGIAPGRGCIQKNGEKHRIVVPVY--- 485
| | | | :
QY 185 FWFECFLRP---RARPHSLMIYSDDLGATW-----HHGRLIK-----PM 221
| | | | :
Db 486 -----YTNKGQSSAVIYSDSGKNWTIGESPNDNRKLKENGKIINTSLSDDAPO 536
| | | | :
QY 222 VTVECEVAEYIGKAGHPVLYCSARTNNRHRAEALSIDHGECKFKPVLSHOLCEPHGCQG 281
| | | | :
Db 537 LT-EQVVEN--PNGQLKLF--NRNLGYLINATSFdGATWDETVERKDTNVLEPY-CQL 590
| | | | :
QY 282 SVVSFCPLEIPGCCQDLAGEDAPAIOQSPILCSSVRPEAGTLESWLLSYSHPTNKKR- 340
| | | | :
Db 591 SVINY-----SQKVDGRDA-----VFISNPNARSRS 616
| | | | :
QY 341 ----RVLDL----GYLNOSPLEAACWSRWLHCGPCGYSDLAALNEGLFGCLFECCGK 392
| | | | :
Db 617 NGTVRICLINQVYENGEPKYEFDWMKNLKPGYYAYSCLTSLSN-GNIGLLYEGTFS 675
| | | | :
QY 393 QECEQIAFRL 402
| | | | :
Db 676 EENSIEIENL 685
| | | | :

RESULT 4
Q4562 PRELIMINARY; PRT; 901 AA.

ID Q4562 PRELIMINARY; PRT; 901 AA.
AC Q4562;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIALLIDASE.
GN NANH.
OS Actinomyces viscosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Actinomycineae; Actinomycetaceae; Actinomycetes.
OX NCBI_TaxID=1656;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T14V;
RX MEDLINE=93114861; PubMed=8418033;
RA Yeung M.K.;
RT "Complete nucleotide sequence of the Actinomyces viscosus T14V
RT sialidase gene; Presence of a conserved repeating sequence among
RT strains of Actinomyces sp.";
RL Infect. Immun. 61:109-116(1993).
DR EMBL; L06898; AAA1932.1; -;
DR HSSP; Q02834; IEUR.
DR InterPro; IPR002860; BNR.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 22.42 seconds
(without alignments)
466.287 Million cell updates/sec

Title: US-09-820-155-2
Perfect score: 2348
Sequence: 1 MEEVTSFSSPLFQOEDKR.....LSHVQDCSTPGMNSPKK 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2348	100.0	428	4	US-09-423-340-2
2	1916.5	81.6	428	4	US-09-423-340-4
3	692	29.5	379	2	US-08-871-074-26
4	201.5	8.6	626	6	5268290-2
5	110	4.7	51	2	US-08-871-074-11
6	110	4.7	51	5	PCT-US94-05471-11
7	98	4.2	931	3	US-08-936-135-16
8	92	3.9	909	3	US-08-936-135-8
9	91.5	3.9	782	2	US-09-146-283-4
10	91.5	3.9	782	3	US-08-579-823A-4
11	91.5	3.9	782	4	US-09-344-195-4
12	91.5	3.9	926	3	US-08-936-135-14
13	87.5	3.7	901	3	US-08-936-135-22
14	87.5	3.7	914	3	US-08-936-135-12
15	87	3.7	906	3	US-08-936-135-24
16	86.5	3.7	1523	4	US-09-182-024-2
17	86.5	3.7	2471	1	US-08-185-432-16
18	86.5	3.7	2471	1	US-08-083-590A-19
19	86.5	3.7	2471	3	US-08-532-384-19
20	85	3.6	794	1	US-07-885-972A-2
21	85	3.6	794	1	US-07-885-972A-4
22	85	3.6	794	2	US-08-865-203-2
23	85	3.6	794	2	US-08-745-880-2
24	85	3.6	794	2	US-08-745-880-4
25	85	3.6	794	2	US-08-480-382-2
26	85	3.6	794	2	US-08-480-382-4
27	85	3.6	794	2	US-07-849-420-2

28	85	3.6	794	4	US-09-253-854-2	Sequence 2, Appli
29	85	3.6	794	4	US-08-955-424-2	Sequence 2, Appli
30	85	3.6	909	3	US-08-936-135-18	Sequence 18, Appl
31	84.5	3.6	926	3	US-08-936-135-20	Sequence 20, Appl
32	84.5	3.6	1255	2	US-08-623-101-2	Sequence 2, Appli
33	84.5	3.6	1255	2	US-08-356-786-2	Sequence 9, Appli
34	84	3.6	484	2	US-08-252-493C-9	Sequence 9, Appli
35	84	3.6	484	3	US-09-276-197-9	Sequence 10, Appl
36	84	3.6	909	3	US-08-936-135-10	Sequence 6, Appli
37	83.5	3.6	1049	4	US-09-522-666-6	Sequence 8, Appli
38	83.5	3.6	1063	1	US-08-127-499A-8	Sequence 8, Appli
39	83.5	3.6	1063	1	US-08-482-847-8	Sequence 54, Appl
40	83.5	3.6	1141	1	US-08-131-365B-54	Sequence 54, Appl
41	83.5	3.6	1141	2	US-08-668-123-54	Sequence 4, Appli
42	83.5	3.6	2629	2	US-08-751-189-4	Sequence 4, Appli
43	83.5	3.6	2629	2	US-09-060-836-4	Sequence 4, Appli
44	83.5	3.6	2629	4	US-09-184-445-4	Sequence 4, Appli
45	82.5	3.5	1155	1	US-08-286-889-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-09-423-340-2
; Sequence 2, Application US/09423340
; Patent No. 6225454
; GENERAL INFORMATION:
; APPLICANT: MITAGI, Taeko
; APPLICANT: WADA, Tadashi
; APPLICANT: YOSHIKAWA, Yuko
; TITLE OF INVENTION: STALIDASE LOCALIZED IN PLASMA MEMBRANE AND
; TITLE OF INVENTION: DNA CODING FOR THE SAME
; FILE REFERENCE: 0P699
; CURRENT APPLICATION NUMBER: US/09/423,340
; EARLIER FILING DATE: 1999-11-22
; EARLIER APPLICATION NUMBER: JP 9-132174
; EARLIER FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Bos primigenius
US-09-423-340-2

Query Match	100.0%	Score	2348	DB	4	Length	428
Best Local Similarity	100.0%	Pred. No.	4.2e-251				
Matches	428	Conservative	0	Mismatches	0	Indels	0
Qy	1	MEEVTSFSSPLFQOEDKRGVYTRIPALIIYVPPAHTFLAFAEKRSSKDEDAHLHLRR	60				
Db	1	MEEVTSFSSPLFQOEDKRGVYTRIPALIIYVPPAHTFLAFAEKRSSKDEDAHLHLRR	60				
Qy	61	GLRTGQSVQWVPLSLMKATLPGRHTMPCPWERKSGYVYLFICVGHVTERQOIMSG	120				
Db	61	GLRTGQSVQWVPLSLMKATLPGRHTMPCPWERKSGYVYLFICVGHVTERQOIMSG	120				
Qy	121	RNPALRCFCISQDAGYSWSDVRLTEEVIGPEVTHWATFVGPCHGHIQLOSGRLIIPAYA	180				
Db	121	RNPALRCFCISQDAGYSWSDVRLTEEVIGPEVTHWATFVGPCHGHIQLOSGRLIIPAYA	180				
Qy	181	YYIFWFFCFRLPYRARPMSLMYSDDLGATWHHGRLLKPMVTVECEVAEIGKAGHPVL	240				
Db	181	YYIFWFFCFRLPYRARPMSLMYSDDLGATWHHGRLLKPMVTVECEVAEIGKAGHPVL	240				
Qy	241	YCSARTPNRRAEALSIDHGEFCFOKPVLSHOLCEPHCCQGSWSFCPLEIPGGCQDLAG	300				
Db	241	YCSARTPNRRAEALSIDHGEFCFOKPVLSHOLCEPHCCQGSWSFCPLEIPGGCQDLAG	300				
Qy	301	EDAPAIQCSPLLCSSVRPEPEAGTLESWLLYSHTPNKRRVLDLGIYLNQSPLEAACWSR	360				
Db	301	EDAPAIQCSPLLCSSVRPEPEAGTLESWLLYSHTPNKRRVLDLGIYLNQSPLEAACWSR	360				

Db 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSESLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360
QY 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFTDRILSHVQDCSTPG 420
Db 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFTDRILSHVQDCSTPG 420
QY 421 MNSEPSKK 428
Db 421 MNSEPSKK 428

RESULT 2
US-09-423-340-4
; Sequence 4, Application US/09423340
; Patent No. 6223454
; GENERAL INFORMATION:
; APPLICANT: MIYAGI, Taeko
; APPLICANT: WADA, Tadashi
; APPLICANT: YOSHIKAWA, YUKO
; TITLE OF INVENTION: SIALIDASE LOCALIZED IN PLASMA MEMBRANE AND
; TITLE OF INVENTION: DNA CODING FOR THE SAME
; FILE REFERENCE: OP699
; CURRENT APPLICATION NUMBER: US/09/423,340
; CURRENT FILING DATE: 1999-11-22
; EARLIER APPLICATION NUMBER: JP 9-132174
; EARLIER FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-340-4

Query Match 81.6%; Score 1916.5; DB 4; Length 428;
Best Local Similarity 82.4%; Pred. No. 2.2e-203;
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;

QY 1 MEEVTCSSPFLQOQDKRGVYRIPALIVVPPHTFLFAEKRSSKDEDALHLVLR 60
Db 1 MEEVTCSSPFLQOQDKRGVYRIPALIVVPPHTFLFAEKRSSKDEDALHLVLR 60

QY 61 GLRTQGVQWNEPLKSLMKATLPGHRTMPCVWERKSGYVYLFICVQGHVTERQQTMSG 120
Db 61 GLRTQGVQWNEPLKSLMKATLPGHRTMPCVWERKSGYVYLFICVQGHVTERQQTMSG 120

QY 121 RNPARLCFICSDAGYSWSVDROLTEEVIGPEVTHWATFVAGPGHGIQLOSGLRIIPAYA 180
Db 121 RNPARLCFICSDAGYSWSVDROLTEEVIGPEVTHWATFVAGPGHGIQLOSGLRIIPAYA 180

QY 181 YVTPWFECFRLPYRARPHSLMIYSDDLGTATWHHGLRIKPMVTVECEVAEVIKAGHPVL 240
Db 181 YVTPWFECFRLPYRARPHSLMIYSDDLGTATWHHGLRIKPMVTVECEVAEVIKAGHPVL 240

QY 241 YCSARTPNRRHRAEALSDHGECKQPKVLSHQLCEPHPGCGSVVSCFPLIPGCGQDLAG 300
Db 241 YCSARTPNRRHRAEALSDHGECKQPKVLSHQLCEPHPGCGSVVSCFPLIPGCGQDLAG 300

QY 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSESLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360
Db 301 EDAPAIQSSPLLCSSVRPEPEAGTSLSESLLYSHPTNKKRRVDLGIYLNQSPLEAACWSR 360

QY 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFTDRILSHVQDCSTPG 420
Db 361 PWILHCGPCGYSDLAALENGLFCGCTKQCEQIAFRFTDRILSHVQDCSTPG 420

QY 421 MNSEPSK 427
Db 421 MNSEPSK 427

RESULT 3

US-08-871-074-26
; Sequence 26, Application US/08871074
; Patent No. 5928915
; GENERAL INFORMATION:
; APPLICANT: Warner, Thomas C.
; APPLICANT: Sliwowski, Mary B.
; TITLE OF INVENTION: CHO Cell Sialidase By Recombinant DNA
; TITLE OF INVENTION: Technology
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,074
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/711,373
; FILING DATE:
; APPLICATION NUMBER: US/08/396,367
; FILING DATE: 28-FEB-1995
; APPLICATION NUMBER: 08/187327
; FILING DATE: 25-JAN-1994
; APPLICATION NUMBER: 08/062586
; FILING DATE: 17-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: 826P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-871-074-26

Query Match 29.5%; Score 692; DB 2; Length 379;
Best Local Similarity 38.6%; Pred. No. 5.1e-68;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;

QY 13 LFQEDKRGVYRIPALIVVPPHTFLFAEKRSSKDEDALHLVLRRLTQGS--VOW 70
Db 13 LFQEDKRGVYRIPALIVVPPHTFLFAEKRSSKDEDALHLVLRRLTQGS--VOW 70

QY 71 EPLKSLMKATLPGHRTMPCVWERKSGYVYLFICVQGHVTERQQTMSGRNPARLCFIC 130
Db 71 EPLKSLMKATLPGHRTMPCVWERKSGYVYLFICVQGHVTERQQTMSGRNPARLCFIC 130

QY 69 QAEVVTQAYLEGRHSMSPCPDYDKQRTLFLFFIAVRGOISHHQLOQTGVNTRLCIT 128
Db 69 QAEVVTQAYLEGRHSMSPCPDYDKQRTLFLFFIAVRGOISHHQLOQTGVNTRLCIT 128

QY 131 SODAGYSWSVDROLTEEVIGPEVTHWATFVAGPGHGIQLO--SGRLIIPAYAYIPFWFF 188
Db 131 SODAGYSWSVDROLTEEVIGPEVTHWATFVAGPGHGIQLO--SGRLIIPAYAYIPFWFF 188

QY 129 STDHGKTSWAVQDLDTTIGSTHODWATFVAGPGHGIQLO--SGRLIIPAYAYIPFWFF 182
Db 129 STDHGKTSWAVQDLDTTIGSTHODWATFVAGPGHGIQLO--SGRLIIPAYAYIPFWFF 182

QY 189 CFRPL--YRAPHSLMIYSDDLGTATWHHGLRIKPMVTVECEVAEVIKAGHPVL 246
Db 189 CFRPL--YRAPHSLMIYSDDLGTATWHHGLRIKPMVTVECEVAEVIKAGHPVL 246

QY 183 --KQPIHAPAPSAFCFLSHDHGSGTWELGHFVS--QNSLEQCAVEV--GTGAERVVYLNARS 238
Db 183 --KQPIHAPAPSAFCFLSHDHGSGTWELGHFVS--QNSLEQCAVEV--GTGAERVVYLNARS 238

QY 247 PNRHRAEALSDHGECKQPKVLSHQLCEPHPGCGSVVSCFPLIPGCGQDLAGDAPAI 306
Db 247 PNRHRAEALSDHGECKQPKVLSHQLCEPHPGCGSVVSCFPLIPGCGQDLAGDAPAI 306

QY 239 CLGARVQAOQSPNSGLDFQDNQVYVSKLVEPPKCGHGVIAF----- 278
Db 239 CLGARVQAOQSPNSGLDFQDNQVYVSKLVEPPKCGHGVIAF----- 278

```

RESULT      5
US-08-871-074-11
; Sequence 11, Application US/08871074
; Patent No. 5928915
; GENERAL INFORMATION:
; APPLICANT: Warner, Thomas C.
; APPLICANT: Slikowski, Mary B.
; TITLE OF INVENTION: CHO Cell Sialidase By Recombinant DNA
; TITLE OF INVENTION: Technology
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

```

```

RESULT      6
PCT-US94-05471-11
: Sequence 11, Application PC/TUS9405471
: GENERAL INFORMATION:
:
: APPLICANT: Genentech, Inc.
: TITLE OF INVENTION: CHO Cell Stialdase By Recombinant DNA Technology
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/05471
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/062586

```


Db 756 HGRILPSYDMEYQIVREGVIGKG-----I: :||| :
QY 272 LCEPHPCGQSVVFCPLPIPGCCQDLAGEDAPAIQQSPLLCSSVRPEAGTILSESLL 331
Db 793 TDVPLENCEPIAFAD-EYEGDWSNSSSTSGA-----GDPSGK-EKSWLY 838
QY 332 YSHP--TNKRRVDLGIYLNQSPLEAACWSRPWTLHCGPCGYSLAA-----LEN 379
Db 839 TLDPLITIAMSLSGLVLLG-----ATC--AGLLIYC-TCSYGLSSRSCTTLEN 885

RESULT 9
US-09-146-283-4
: Sequence 4, Application US/09146283
: Patent No. 5976546
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Compositions
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/146,283
: FILING DATE: 03-SEPT-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 782 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 3.9%; Score 91.5; DB 2; Length 782;
Best Local Similarity 20.7%; Pred. No. 0.41;
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;
QY 94 ERKSGYVLF--ICVQGHVTERQQIMSGR---NPARLCFICSDAGYSVDVRLTTEE 147
Db 404 EETGYLYISAWPDSLPDLVSFQNLQVIRGLHNGA--YSLTQGLGISWLGRLSREL 461
QY 148 VIGPEVTHWATFVAGPGHGIQSGRLIIPAYAYIIPWFECFRLPYRAPHSLMIYSD 207
Db 462 GSGALAIH-----HNTHL-----CFVHTVP-WDQLFRNPHOALLHTANRPEDE 503
QY 208 L-----GATWHHGRLLKPMWTVCE-----VAEVIKAGHPVLYCSAR- 245
Db 504 CVGEGLAGHQLCARGHCWGP-----PTQVCNCSQFLRGQECVEECRVLQGLPREYVNAH 559

QY 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271
Db 560 CLPCHPCQPNQSVTCFGEADOCVACAHYKDPFVCARCPSCVKPDLSLMPWKFPDE 619
QY 272 --LCEPHPCGQSVVFCPLPIPGCCQDLAGEDAPAIQQSPLLCSSVRPEAGTILSES 329
Db 620 EGACQP-----CPINCTHSCVDLDDKGCPCQASPLTSLAPARSPSPSTQPW 668
RESULT 10
US-08-579-823A-4
: Sequence 4, Application US/08579823A
: Patent No. 6080409
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 782 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match 3.9%; Score 91.5; DB 3; Length 782;
Best Local Similarity 20.7%; Pred. No. 0.41;
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;
QY 94 ERKSGYVLF--ICVQGHVTERQQIMSGR---NPARLCFICSDAGYSVDVRLTTEE 147
Db 404 EETGYLYISAWPDSLPDLVSFQNLQVIRGLHNGA--YSLTQGLGISWLGRLSREL 461
QY 148 VIGPEVTHWATFVAGPGHGIQSGRLIIPAYAYIIPWFECFRLPYRAPHSLMIYSD 207
Db 462 GSGALAIH-----HNTHL-----CFVHTVP-WDQLFRNPHOALLHTANRPEDE 503
QY 208 L-----GATWHHGRLLKPMWTVCE-----VAEVIKAGHPVLYCSAR- 245
Db 504 CVGEGLAGHQLCARGHCWGP-----PTQVCNCSQFLRGQECVEECRVLQGLPREYVNAH 559
QY 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271

Db 560 CLPCHPECQONGSVTCFGPEADQCVACAHYKDPFCVARCPGKVPDLISYMPDIWKFPDE 619
QY 272 --LCEPHPGCGQSVWFCPLIEPGGQDLAGEDAPAIQQSPILCSSLVRPEAGTISEW 329
Db 620 EGACQP-----CPINCTHSCVDLDDKCPAEQRAASPLTSLEAPARSPSPSTQPW 668
RESULT 11
US-09-344-195-4
; Sequence 4, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; Ruegg, Curtis L.
; Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-344-195-4
Query Match 3.9%; Score 91.5; DB 4; Length 782;
Best Local Similarity 20.7%; Pred. No. 0.41;
Matches 62; Conservative 27; Mismatches 112; Indels 99; Gaps 13;
QY 94 ERKSGVYLFF---ICVQGHVTERQIMSGR---NPARLCFCISQDAGYSWSDVRLTTEE 147
Db 404 EEITGYLISANPDLISVFNQNLQVIRGRLHNGA--YSLTQGLGSLGLRLSLREL 461
QY 148 VIGPEYTHWATFVGPHGIGLOSGRLIIPAYAYIIPFWFFCFRLPYRAPHSLMTYSDD 207
Db 462 GSGLALIH-----HNTHL-----CFVHTVP-WDQLFRNPHQALLHTANRPEDE 503
QY 208 L-----GATWHGRLIKPMVTVECE-----VAEVIGKAGHPVLYCSAR- 245
Db 504 CVGEGLACHQLCARGHCWGP-----PTQCVCNSQFLRGQECVEECRVLQGLPREYVNAH 559

QY 246 -----TPNRHRAEALSIDHGECFQ-----KPVLSHQ----- 271
Db 560 CLPCHPECQONGSVTCFGPEADQCVACAHYKDPFCVARCPGKVPDLISYMPDIWKFPDE 619
QY 272 --LCEPHPGCGQSVWFCPLIEPGGQDLAGEDAPAIQQSPILCSSLVRPEAGTISEW 329
Db 620 EGACQP-----CPINCTHSCVDLDDKCPAEQRAASPLTSLEAPARSPSPSTQPW 668
RESULT 12
US-08-936-135-14
; Sequence 14, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-936-135-14
Query Match 3.9%; Score 91.5; DB 3; Length 926;
Best Local Similarity 22.9%; Pred. No. 0.54;
Matches 55; Conservative 30; Mismatches 90; Indels 65; Gaps 13;
QY 170 QSGRLIIPAYAYIIPFWFFCFRLPYRA-----RPHSLMTYSDDLGTATH 213
Db 698 QGRLISP--PVHLPRSPVCMFEFYQAMGHHGVALQVREASQESKLLWVIREDDQSEWK 755
QY 214 HGRLIKPMVTVECEVA--EVIKAGHPVLYCSARTPNRHRAEALSIDHGECFQKPVLSHQ 271
Db 756 HGRILPSYDMEYQIVFEGVIGKG-----RSGEISGDD-----IRIS 792
QY 272 LCEPHPGCGQSVVFCPLIEP-----GGQDLAGEDAPAIQQSPILCSSLVRPEAGTIS 326
Db 793 TDVPLENCMEPIISAFV-DIPETHGGEYEDIDDEYEGDWSNSSSTSGAGDPSSGK-E 850
QY 327 ESKWLLYSHP--INKKRRVDLGIYLNQSPLEAACWSRPWILHCGPCGYSDLAA-----LEN 379
Db 851 KSWLYTLDILITITIAMSSGLVLLG-----ATC--AGLLLYC--TCSYSLGSSRSCTTLEN 902

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 211.59 Seconds
(without alignments)
711.979 Million cell updates/sec

Title: US-09-820-155-2
Perfect score: 2348
Sequence: 1 MEEVTSCTSFSSPLFQEDKR.....LSHVQGDCTPGMNSPESKK 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2348	100.0	428	22	US-09-820-155-2
2	1916.5	81.6	428	22	US-09-820-155-4
3	875	37.3	161	22	US-09-820-155-25
4	653	27.8	372	19	US-09-556-870A-2
5	520	22.1	362	26	US-60-207-359-294
6	476.5	20.3	240	22	US-09-820-155-26
7	295	12.6	245	1	PCT-US01-01240-372

8	295	12.6	245	1	PCT-US01-01324-1918	Sequence 1918, Ap
9	295	12.6	245	21	US-09-764-872-372	Sequence 372, App
10	295	12.6	245	21	US-10-079-379-1918	Sequence 1918, Ap
11	255.5	10.9	425	21	US-09-758-442-464	Sequence 464, App
12	213	9.1	151	1	PCT-US01-01354-18118	Sequence 18118, A
13	213	9.1	151	21	US-09-764-905-18118	Sequence 18118, A
14	213	9.1	151	24	US-10-092-399-18118	Sequence 18118, A
15	177	7.5	413	18	US-09-417-507-38401	Sequence 38401, A
16	147.5	6.3	937	19	US-09-583-110-3166	Sequence 3166, Ap
17	147.5	6.3	975	15	US-09-107-433-3032	Sequence 3032, Ap
18	140.5	6.0	683	24	US-10-091-007-64	Sequence 64, Appl
19	138	5.9	248	1	PCT-US00-05883-958	Sequence 958, App
20	138	5.9	248	23	US-09-925-299-958	Sequence 958, App
21	137.5	5.9	394	21	US-09-738-626-5219	Sequence 5219, Ap
22	110	4.7	51	12	US-08-871-076-5	Sequence 9, Appli
23	108.5	4.6	642	13	US-08-911-393-4	Sequence 4, Appli
24	108.5	4.6	642	23	US-09-955-909-4	Sequence 4, Appli
25	106.5	4.5	669	24	US-10-086-913-2	Sequence 2, Appli
26	106.5	4.5	1060	13	US-08-911-393-2	Sequence 2, Appli
27	106.5	4.5	1060	23	US-09-955-909-2	Sequence 2, Appli
28	105.5	4.5	666	21	US-09-745-008-2	Sequence 2, Appli
29	104.5	4.5	763	1	PCT-US01-08631-3127	Sequence 3127, A
30	102	4.3	567	18	US-09-488-725A-1803	Sequence 1803, Ap
31	101	4.3	571	21	US-09-760-466-907	Sequence 907, App
32	98	4.2	931	1	PCT-US98-14290-16	Sequence 16, Appl
33	98	4.2	931	18	US-09-439-711-16	Sequence 16, Appl
34	97	4.1	583	21	US-09-760-466-1379	Sequence 1379, Ap
35	97	4.1	925	15	US-09-116-473-2	Sequence 2, Appli
36	96	4.1	347	26	US-60-242-679-1068	Sequence 1068, Ap
37	96	4.1	391	24	US-10-015-127-10231	Sequence 10231, A
38	94.5	4.0	596	18	US-09-488-725A-5375	Sequence 5375, Ap
39	94.5	4.0	2820	20	US-09-614-150-16680	Sequence 16680, A
40	94	4.0	820	26	US-60-150-584-690	Sequence 690, App
41	93	4.0	1078	26	US-60-207-315-428	Sequence 428, App
42	93	4.0	1078	26	US-60-212-659-523	Sequence 523, App
43	93	4.0	1078	26	US-60-230-435-1751	Sequence 1751, Ap
44	92	3.9	476	24	US-10-029-386-32293	Sequence 32293, A
45	92	3.9	479	1	PCT-US01-07143-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-09-820-155-2
; Sequence 2, Application US/09820155
; GENERAL INFORMATION:
; APPLICANT: Miyagi, Taeko
; APPLICANT: Wada, Tadashi
; APPLICANT: Yoshikawa, Yuko
; TITLE OF INVENTION: Stalidase Localized in the Plasma Membrane
; FILE REFERENCE: 2011-0046 (45455-252923)
; CURRENT APPLICATION NUMBER: US/09/820,155
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/423,340
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/JP98/02072
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Bos primigenius taurus
US-09-820-155-2

Query Match 100.0%; Score 2348; DB 22; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.6e-217;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEEVTSCTSFSSPLFQEDKRGVYRIPALTYVPATLAFAEKRSKDEDAHLVLR 60
|||||

Db 1 MEEVTSFSSPLFOQEDKRGVYTRIPALIVYPAPHTLAFAPKRSKDEADLHLVLR 60
Qy 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIMSG 120
Db 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIMSG 120
Qy 121 RNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180
Db 121 RNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180
Qy 181 YIPWFVFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVCEVAEVTGAGHPVL 240
Db 181 YIPWFVFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVCEVAEVTGAGHPVL 240
Qy 241 YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300
Db 241 YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300
Qy 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHPTNKKRRVLDGIYLNQSPLEAACWSR 360
Db 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHPTNKKRRVLDGIYLNQSPLEAACWSR 360
Qy 361 PWILHCGPCGYSDLAALNEGLFGLFCGCKQCEQIAFRLFTDREILSHVQDCSTPG 420
Db 361 PWILHCGPCGYSDLAALNEGLFGLFCGCKQCEQIAFRLFTDREILSHVQDCSTPG 420
Qy 421 MNSEPSKK 428
Db 421 MNSEPSKK 428

RESULT 2

US-09-820-155-4
; Sequence 4, Application US/09820155
; GENERAL INFORMATION:
; APPLICANT: Miyagi, Taeko
; APPLICANT: Wada, Tadashi
; APPLICANT: Yoshikawa, Yuko
; TITLE OF INVENTION: Sialidase Localized in the Plasma Membrane
; FILE REFERENCE: 20111-0046 (45455-252923)
; CURRENT APPLICATION NUMBER: US/09/820,155
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/423,340
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/JP98/02072
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-820-155-4

Query Match 81.6%; Score 1916.5; DB 22; Length 428;
Best Local Similarity 82.4%; Pred. No. 8e-176;
Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;

Qy 1 MEEVTSFSSPLFOQEDKRGVYTRIPALIVYPAPHTLAFAPKRSKDEADLHLVLR 60
Db 1 MEEVTSFSSPLFOQEDKRGVYTRIPALIVYPAPHTLAFAPKRSKDEADLHLVLR 60
Qy 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIMSG 120
Db 61 GLRTGQSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQGHVTERQIMSG 120
Qy 121 RNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180
Db 121 RNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQLOSGRLIIPAYA 180
Qy 181 YIPWFVFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVCEVAEVTGAGHPVL 240
Db 181 YIPWFVFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVCEVAEVTGAGHPVL 240

Db 181 YIPWFVFCFRLPYRAPHSLMIYSDDLGATWHHGRLLKPMVTVCEVAEVTGAGHPVL 240
Qy 241 YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300
Db 241 YCSARTPNRRAEALSIDHGECFOKPVLSHOLCEPHCGQSVVSFCPLEIPGGCQDLAG 300
Qy 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHPTNKKRRVLDGIYLNQSPLEAACWSR 360
Db 301 EDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHPTNKKRRVLDGIYLNQSPLEAACWSR 360
Qy 361 PWILHCGPCGYSDLAALNEGLFGLFCGCKQCEQIAFRLFTDREILSHVQDCSTPG 420
Db 361 PWILHCGPCGYSDLAALNEGLFGLFCGCKQCEQIAFRLFTDREILSHVQDCSTPG 420
Qy 421 MNSEPSK 427
Db 420 RN--PSQ 424

RESULT 3

US-09-820-155-25
; Sequence 25, Application US/09820155
; GENERAL INFORMATION:
; APPLICANT: Miyagi, Taeko
; APPLICANT: Wada, Tadashi
; APPLICANT: Yoshikawa, Yuko
; TITLE OF INVENTION: Sialidase Localized in the Plasma Membrane
; FILE REFERENCE: 20111-0046 (45455-252923)
; CURRENT APPLICATION NUMBER: US/09/820,155
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/423,340
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/JP98/02072
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-820-155-25

Query Match 37.3%; Score 875; DB 22; Length 161;
Best Local Similarity 98.8%; Pred. No. 1e-75;
Matches 159; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 49 KDEDALHLVLRRLTGOSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQ 108
Db 1 KDEDALHLVLRRLTGOSVQWPEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFTCVQ 60
Qy 109 GHVTERQIMSGRNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQ 168
Db 61 GHVTERQIMSGRNPARLCFICSODAGYSWSDVRLTEEVIGPEVTHWATFVAVGPGHGIQ 120
Qy 169 LOSGRLIIPAYAYIPWFVFCFRLPYRAPHSLMIYSDDLG 209
Db 121 LOSGRLIIPAYAYIPWFVFCFRLPYRAPHSLMIYSDDLG 161

RESULT 4

US-09-556-870A-2
; Sequence 2, Application US/09556870A
; GENERAL INFORMATION:
; APPLICANT: Yu, Robert
; APPLICANT: Chris, Fronda
; APPLICANT: Zeng, Guichao
; TITLE OF INVENTION: CONA SEQUENCE OF MOUSE BRAIN SIALIDASE GENE
; FILE REFERENCE: SEQUENCE (02940110AA)
; CURRENT APPLICATION NUMBER: US/09/556,870A
; CURRENT FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: US 60/130,712
; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-556-870A-2

Query Match 27.8%; Score 653; DB 19; Length 372;
Best Local Similarity 39.6%; Pred. No. 9.1e-54;
Matches 157; Conservative 59; Mismatches 114; Indels 66; Gaps 15;

QY 1 MBEVTSFSSFLFOED--KRGV-TYRIPALIYPPAHTFLAFAEKRSSKDEDAHLV 57
DB 1 MEDLRPMA-TCPVLQKETLFRGVHAYRIPALLYLKKQKTLAFAERRASKTDEHAELIV 59

QY 58 LRRLG--RTGQSVQWELPLSKMATLPGHRTMNPVWERKSGYVYLFFICVQGHVTERQ 115
DB 60 LRSGSYNEATNRVKWQPEBVTQAOLGHRSMNCPCLMTSKKD-LFLFFIAVPGRVSEHH 118

QY 116 QTMGRNPARLFCFICSQDAGYSWSDVRDLTEEVIGPEVTHWATFVPGCHGILQ--SGR 173
DB 119 QLHTKVNVTRLCCVSDTGRMTSPQDLTETTIGSTHOEWATEFVPGHCLQLRNPAGS 178

QY 174 LIIPAYAY--IPWFECFRLPYRAPHSLMIYSDDLGATWHHGRLIKPMVTVE 225
DB 179 LLVPAYAYKRLHPAQKPTFF-AFCF-----ISLDHGTWKLVGNFAEN-SUE 223

QY 226 CEVAEYIGKAGHPVLYCSARTPNRRAEALSIDHGECFQKPVLSHOLCBPPHGCQGSVYS 285
DB 224 CQVAEV-GTGAQRMVYLNARSFLGARVQAQSPNDGLDFODNRVSKLVEPPHGGCHSVVA 282

QY 286 FCPLEIPGCCQDLAGEDAPAIQOSPLLCSSVRPEPAGTLSWSWLLYSHPTNKKRVDLG 345
DB 283 F-----HNPT-----SKPHA---LDTWLLYTHPTDSRNRNLG 312

QY 346 IYLNQSPLEAACWSRPWLHCPCGYSGLAALENEG 381
DB 313 VYLNQMPDLPTAWSEPTLLAMGICAYSD---LQNMG 345

RESULT 5
US-60-207-359-294
; Sequence 294; Application US/60207359
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000592
; CURRENT APPLICATION NUMBER: US/60/207,359
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 362
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-359-294

Query Match 22.1%; Score 520; DB 26; Length 362;
Best Local Similarity 35.5%; Pred. No. 6.3e-41;
Matches 117; Conservative 42; Mismatches 109; Indels 62; Gaps 6;

QY 154 THWATFVAGPGHGIQSGRLIIPAYAYIPWFECFRLPYRAPHSLMIYSDDLGATWH 213
DB 30 SYWATFVAGPGHGVQLPSGRLVLPATYRVD-RRECFGKICRTSPHSFAFYSDDHGRWTR 88

QY 214 HGRLIKPMVTVECEVAEYIGKAGHPVLYCSARTPNRRAEALSIDHGECFQKPVLSHQZ 273
DB 89 CGGLVPLNRSGEQLAADVGGAGSFLYCNARSPLGSRVQALSTDBGTSTFLPAERVASLP 148

QY 274 EPPHGCQGSVVSFCPLPIPGGCQD---LAGEDAPAIQOSPLLCSSVRPEAGTL----- 325
DB 149 ETAWCCQGSIVGF-PAPAPNRPRDDSSWSVGPGL--QPPLLGPGVHPPPEAAVDPGRC 205

QY 326 -----SESWLLYSHPTNKK 339
DB 206 QVPGGPFSLRQPRGPRGPRGVSQDVGSWTLALPMPFAAPQSPPTWLLYSHPVGRR 265

QY 340 RYVDLGIYLNQSPLEAACWSRPWLHCPCGYSGLAAL-----ENEGFLGCLFECGCKQBC 395
DB 266 ALRHGIRLSQSPDLPRSWTEPMWYIEGSPGYSGLASIGPAPEGLVFACLYESGARTSY 325

QY 396 EGIATRLFTDREILSHVGDCTSTPGMNSRP 425
DB 326 DEISFCTSLREVLNVNVPASPPNLDGKP 355

RESULT 6
US-09-820-155-26
; Sequence 26; Application US/09820155
; GENERAL INFORMATION:
; APPLICANT: Miyagi, Taeko
; APPLICANT: Wada, Tadashi
; APPLICANT: Yoshikawa, Yuko
; TITLE OF INVENTION: Sialidase Localized in the Plasma Membrane
; FILE REFERENCE: 20111-0046 (45455-252923)
; CURRENT APPLICATION NUMBER: US/09/820,155
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/423,340
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/JP98/02072
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Rat
US-09-820-155-26

Query Match 20.3%; Score 476.5; DB 22; Length 240;
Best Local Similarity 43.2%; Pred. No. 5.7e-37;
Matches 108; Conservative 40; Mismatches 71; Indels 31; Gaps 8;

QY 12 PLFQOE---DKRGVYRIPALIYPPAHTFLAFAEKRSSKDEDAHLVLRRLGRTGQS- 67
DB 5 PVLQKETLFTHTVEYAYRIPALLYLKKQKTLAFAEKRASRTDEHAELIVLRGRSYNGATN 64

QY 68 -VQWPEPLSKMATLPGHRTMNPVWERKSGYVYLFFICVQGHVTERQIAMSGRNPARL 126
DB 65 HVKQWPEEYVTOAQLEGRHSMNCPCLYDKQTKTLFFFAVPGRVSEHQLOTRVNVTRL 124

QY 127 CFICSQDAGYSWSDVRDLTEEVIGPEVTHWATFVPGCHGILQ--QSGRLIIPAYAY-- 182
DB 125 CRVTSTDYGNWSPVQDLTETTIGSTHQDWATEFVPGHCLQRNPAQSLLVPAYAYRL 184

QY 183 -----IPWFECFRLPYRAPHSLMIYSDDLGATWHHGRLIKPMVTVECEVAEYIGKAG 236
DB 185 HPVHKPTPE-AFCF-----ISLDHGTWELGNFVSEN-SLSCQVAEV-GTGA 228

QY 237 HPVLYCSART 246
DB 229 HRVYVLNARS 238

RESULT 7
PCT-US01-01240-372
; Sequence 372; Application PC/TUS0101240
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

Matches	40;	Conservative	8;	Mismatches	9;	Indels	0;	Gaps	0;
Qy	1	MEEVTSFCSPLFQEQDKRGVYTRIPALYYPPAHFTFLFAEKRSSKDEDAHLV	57	: : :	:				
Db	76	MEEVTSFCSPLFQEQDGRGTXRIPALLYIXPTXFLFAEKRSTRDEDAKLX	132	: : :	:				
RESULT 13									
US-09-764-905-18118									
: Sequence 18118, Application US/09764905									
: GENERAL INFORMATION:									
: APPLICANT: Rosen et al.									
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies									
: FILE REFERENCE: PC004									
: CURRENT APPLICATION NUMBER: US/09/764,905									
: CURRENT FILING DATE: 2001-01-17									
: PRIOR APPLICATION NUMBER: 60/179,065									
: PRIOR FILING DATE: 2000-01-31									
: PRIOR APPLICATION NUMBER: 60/180,628									
: PRIOR FILING DATE: 2000-02-04									
: PRIOR APPLICATION NUMBER: 60/214,886									
: PRIOR FILING DATE: 2000-06-28									
: PRIOR APPLICATION NUMBER: 60/217,487									
: PRIOR FILING DATE: 2000-07-11									
: PRIOR APPLICATION NUMBER: 60/225,758									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/220,963									
: PRIOR FILING DATE: 2000-07-26									
: PRIOR APPLICATION NUMBER: 60/217,496									
: PRIOR FILING DATE: 2000-07-11									
: PRIOR APPLICATION NUMBER: 60/225,447									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/218,290									
: PRIOR FILING DATE: 2000-07-14									
: PRIOR APPLICATION NUMBER: 60/225,757									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/226,868									
: PRIOR FILING DATE: 2000-08-22									
: PRIOR APPLICATION NUMBER: 60/216,647									
: PRIOR FILING DATE: 2000-07-07									
: PRIOR APPLICATION NUMBER: 60/225,267									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/216,880									
: PRIOR FILING DATE: 2000-07-07									
: PRIOR APPLICATION NUMBER: 60/225,270									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/251,869									
: PRIOR FILING DATE: 2000-12-08									
: PRIOR APPLICATION NUMBER: 60/235,834									
: PRIOR FILING DATE: 2000-09-27									
: PRIOR APPLICATION NUMBER: 60/234,274									
: PRIOR FILING DATE: 2000-09-21									
: PRIOR APPLICATION NUMBER: 60/234,223									
: PRIOR FILING DATE: 2000-09-21									
: PRIOR APPLICATION NUMBER: 60/228,924									
: PRIOR FILING DATE: 2000-08-30									
: PRIOR APPLICATION NUMBER: 60/224,518									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/236,369									
: PRIOR FILING DATE: 2000-09-29									
: PRIOR APPLICATION NUMBER: 60/224,519									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/220,964									
: PRIOR FILING DATE: 2000-07-26									
: PRIOR APPLICATION NUMBER: 60/241,809									
: PRIOR FILING DATE: 2000-10-20									
: PRIOR APPLICATION NUMBER: 60/249,299									
: PRIOR FILING DATE: 2000-11-17									
: PRIOR APPLICATION NUMBER: 60/236,327									
: PRIOR FILING DATE: 2000-09-29									
: PRIOR APPLICATION NUMBER: 60/241,795									
: PRIOR FILING DATE: 2000-10-20									
: PRIOR APPLICATION NUMBER: 60/244,617									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/226,681									
: PRIOR FILING DATE: 2000-08-22									
: PRIOR APPLICATION NUMBER: 60/225,759									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,213									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/227,182									
: PRIOR FILING DATE: 2000-08-22									
: PRIOR APPLICATION NUMBER: 60/225,214									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/235,836									
: PRIOR FILING DATE: 2000-09-27									
: PRIOR APPLICATION NUMBER: 60/230,438									
: PRIOR FILING DATE: 2000-09-06									
: PRIOR APPLICATION NUMBER: 60/215,135									
: PRIOR FILING DATE: 2000-06-30									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,266									
: PRIOR FILING DATE: 2000-08-14									
: PRIOR APPLICATION NUMBER: 60/225,268									
: PRIOR FILING DATE: 20									

QY 12 PLFOED-----KRGV--TTRIPALIVPPAHTFLAFAE-KRSSKDEDAHLVLRR 60
 IIII : : : ::IIIII:: IIIII : ::I : :II :
 Db 43 PLFSANMASDPKLTGTGEGHSRIPAVVVTTGR-ILFAEGRHHNODFGDINLVYKR 10

QY 61 GLRTGQS----VQWEPKSLMKATLPGHRTM-NPCPVWERKSGYVYLF----- 104
Db 102 TKTTANNGASPSDWEPLREVGS---GAGTWCNPTPVVD-DDNTIYLFSLWNGATYSQNG 157
QY 105 --ICVQGHVTER-QOIMSGRNPRLCFTCSODAGYSWSVDRDLTEEVIGPEVTHWATFAV 161
Db 158 KDVLPDGTVTRKKIDSTWEGRR--HLYLTESRDDGNTWSKPVDLTKE-LTPD--GWAWDAV 212
QY 162 GPGHGIQLOSGRLIIPA 178
Db 213 GPGNGIRLTGTGELVIPA 229

Search completed: October 7, 2002, 15:04:31
Job time: 353 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2002, 14:58:38 ; Search time 79.21 Seconds
(without alignments)
1490.729 Million cell updates/sec

Title: US-09-820-155-2

Perfect score: 2348

Sequence: 1 MEEVTCSSPFLFOQEDKR.....LSHVQGDCTPGMNSPSSKK 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1003375 seqs, 275889267 residues

Total number of hits satisfying chosen parameters: 1003375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Nev.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1916.5	81.6	461	5	US-09-629-469A-14038
2	1501	63.9	418	5	US-09-791-537-56194
3	887.5	37.8	484	6	US-10-104-047-2240
4	711.5	30.3	380	5	US-09-791-537-27331
5	705	30.0	379	5	US-09-791-537-15522
6	692	29.5	379	5	US-09-791-537-114633
7	255.5	10.9	415	5	US-09-791-537-87692
8	255.5	10.9	425	6	US-10-217-623-464
9	245.5	10.5	409	5	US-09-791-537-1830
10	201.5	8.6	605	5	US-09-791-537-73212
11	201.5	8.6	647	5	US-09-791-537-81425
12	200.5	8.5	365	5	US-09-791-537-73210
13	197	8.4	748	5	US-09-791-537-48863
14	186	7.9	1014	5	US-09-791-537-74548
15	172	7.3	694	5	US-09-791-537-140247
16	169	7.2	901	5	US-09-791-537-34557
17	159.5	6.8	548	5	US-09-540-209B-9518
18	154.5	6.6	544	5	US-09-791-537-145838
19	154	6.6	913	5	US-09-791-537-10376
20	150.5	6.4	404	5	US-09-791-537-46533
21	145.5	6.2	404	5	US-09-791-537-10388
22	142.5	6.1	404	5	US-09-791-537-46515
23	140.5	6.0	404	5	US-09-791-537-46534
24	134.5	5.7	329	5	US-09-540-209B-6320
25	132.5	5.6	679	5	US-09-791-537-104002
26	132.5	5.6	762	5	US-09-791-537-85931

27	128.5	5.5	266	5	US-09-791-537-82976	Sequence 82376, A
28	128	5.5	382	5	US-09-791-537-10387	Sequence 10387, A
29	120	5.1	382	5	US-09-791-537-34298	Sequence 34298, A
30	117.5	5.0	773	5	US-09-791-537-12830	Sequence 12830, A
31	113.5	4.8	642	5	US-09-791-537-66354	Sequence 66354, A
32	111.5	4.7	642	5	US-09-791-537-66353	Sequence 66353, A
33	109.5	4.7	642	5	US-09-791-537-66351	Sequence 66351, A
34	109.5	4.7	964	5	US-09-791-537-13325	Sequence 13325, A
35	108.5	4.6	642	5	US-09-791-537-119053	Sequence 119053, A
36	106.5	4.5	879	5	US-09-791-537-1875	Sequence 1875, Ap
37	106.5	4.5	1060	5	US-09-791-537-13327	Sequence 13327, A
38	105	4.5	757	5	US-09-791-537-87869	Sequence 87869, A
39	102	4.3	680	5	US-09-791-537-28744	Sequence 28744, A
40	101.5	4.3	700	5	US-09-791-537-53751	Sequence 53751, A
41	101	4.3	376	5	US-09-791-537-124536	Sequence 124536, A
42	101	4.3	571	6	US-10-212-083-907	Sequence 907, App
43	100	4.3	382	5	US-09-791-537-31493	Sequence 31493, A
44	100	4.3	556	7	US-60-389-987-2952	Sequence 2952, Ap
45	100	4.3	567	7	US-60-389-987-1410	Sequence 1410, Ap

ALIGNMENTS

RESULT 1

US-09-629-469A-14038

; Sequence 14038, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629,469A

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: JP 1999-300253

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/159,590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/183,322

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 19025

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 14038

; LENGTH: 461

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-629-469A-14038

Query Match 81.6%; Score 1916.5; DB 5; Length 461;

Best Local Similarity 82.4%; Pred. No. 1,le-162;

Matches 352; Conservative 32; Mismatches 40; Indels 3; Gaps 2;

Qy 1 MEEVTCSSPFLFOQEDKRGVYTRIPALYYPPAHTFLAFKRSKSSKDEALHLVLR 60

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 34 MEEVTCSSPFLFOQEDKRGVYTRIPALYYPPAHTFLAFKRSKSSKDEALHLVLR 93


```
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 27331
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-27331

Query Match 30.3%; Score 711.5; DB 5; Length 380;
Best Local Similarity 40.7%; Pred. No. 4.4e-55;
Matches 170; Conservative 59; Mismatches 114; Indels 75; Gaps 16;

QY 10 SSPLFOED--KRGV-TYRIPALIVVPPAHTFLAFAEKRRSSKDEDALHLVLRRLG--LR 63
DB 3 SLPLVKESVFQSGAHAYRIPALILYLPQQSLLAFAEQRAKSKDEHAELIVLRGDDYDAP 62
QY 64 TGSVQWEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFICVQGHVTERQOIMSGRNP 123
DB 63 THQ-VQMQAEVVAQARLDGHRSMNCPPLYDAQGTGLFLEFFIAIPGVTEQOQLQTRANV 121
QY 124 ARLCFTCSQAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQL--QSGRLIIPAYAY 181
DB 122 TRLCQVSTDHGRTWSPRODLTDAALGPAYRWSTEA VPGHCLQNDRA RSLVWPAYAY 181
QY 182 Y-----IPWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEIVG 233
DB 182 RKLUHTQRPIS-AFCF-----LSHDGRTWARGHV-AQDTLECOVAEV-- 224
QY 234 KAGHP-VLYCSARTPNRRAEALSIDHGECFQKPVLSHQICE-PPHGCGSGSVVFCPLLEI 291
DB 225 ETGEQVRVTLNARSHLRARVQAGSTNDGLDFQESQLVKLVPPPOGCGSVISF----- 279
QY 292 PGCQCQDAGEDAPAIQOSPLLCSSVRPEPEAGTILSSS-WLLYSHPTNKKRRVDLGIYLNQ 350
DB 280 -----PSPRSGPGSPAQWLLYTHPTSHWQADLGLAYLNP 313
QY 351 SPLEAACMSRPWLHCGPCGYSDLAAL----ENEGLFGCLFECGTQKECEQIAFRFLT 404
DB 314 RPPAPEANSEPVLLAKGSCAISDLQSMGTGPDGSPFLGCLYEIA---NDYEEIVFLMFT 368

RESULT 5
US-09-791-537-15522
; Sequence 15522, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15522
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-15522

Query Match 30.0%; Score 705; DB 5; Length 379;
Best Local Similarity 39.8%; Pred. No. 1.7e-54;
Matches 164; Conservative 63; Mismatches 117; Indels 68; Gaps 14;
```

```
QY 12 PLFQOED--KRGV-TYRIPALIVVPPAHTFLAFAEKRRSSKDEDALHLVLRRLG--RTGQ 66
DB 5 PVLQETLFTGTGVAHYRIPALILYLLKQKTLAFAEKRAKSTDEHAELIVLRGSGYNEATN 64
QY 67 SYOWEPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFICVQGHVTERQOIMSGRNP 126
DB 65 RVKQOPEEVVTAQLEGRHSMNCPPLYDKQTKTLFFIAVPCRVSEHHQLHTKVNVTRL 124
QY 127 CFICSQDAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQLO--SGRLLIIPAYAY-- 182
DB 125 CCVSSTDHGTWSPRODLTETTGTHQEWATFVAGPGHCLQRLNPAGSLLVPAYAYRKL 184
QY 183 -----IPWFFCFRLPYRAPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEIVG 236
DB 185 HPAQKPTPE-AFCF-----ISLDHGTWKLGNFVAEN-SLECOVAEV-GTCA 228
QY 237 HPLVYCSARTPNRRAEALSIDHGECFQKPVLSHQICEPHGCGSGSVVFCPLLEIPGGCQ 296
DB 229 QRMVYLNARSFLGARVQAQSPNDGLDFQDNRVVSKLVEPPHGHGSGVAF----- 278
QY 297 DIAGEDAPAIQOSPLLCSSVRPEPEAGTILSESWLLYSHPTNKKRRVDLGIYLNQSPLEAA 356
DB 279 -----HNPI-----SKPHA---LDTWLLYTHPTDSNRNTNLGVYLNQMLDPT 318
QY 357 CWSRPWLHCGPCGYSDLAAL----ENEGLFGCLFECGTQKECEQIAFRFLT 404
DB 319 ANSEPTLLANGICAYSDLNQMGQPDGSPQFGCLYESG---NYEELIFLIFT 367

RESULT 6
US-09-791-537-114633
; Sequence 114633, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 114633
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Cricetulus griseus
US-09-791-537-114633

Query Match 29.5%; Score 692; DB 5; Length 379;
Best Local Similarity 38.6%; Pred. No. 2.4e-53;
Matches 159; Conservative 64; Mismatches 133; Indels 56; Gaps 11;
```

```
QY 13 LFOQEDKRGVTRIPALIVVPPAHTFLAFAEKRRSSKDEDALHLVLRRLGRTQGS--VOW 70
DB 12 LFOTGD---YAYRIPALILYLSKQKTLAFAEKRLTKTDEHADLFVLRGSGYNADTHQVQW 68
QY 71 EPLKSLMKATLPGHRTMPCPVWERKSGYVYLFFICVQGHVTERQOIMSGRNP 130
DB 69 QAEVVTVQAYLEGRHSMNCPPLYDKQTKTLFFIAVRCOISHHQLQTVGNVTRLCBIT 128
QY 131 SQDAGYSWSDVRLTEEVIGPEVTHWATFVAGPGHGIQLO--SGRLLIIPAYAYIIFWFF 188
DB 129 STDHGKTSVAVQDLTDTTGSTHQDWATFVAGPGHCLQRLNPAGSLLVPAYAYR----- 182
QY 189 CRLP--YRARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEIVGKAGHPVLYCSART 246
DB 183 --KQPIHAPAPSAFCFLSHDGTWELGHFVS-QNSLECOVAEV-GTGAERVVYLNARS 238
QY 247 PNHRAEALSIDHGECFQKPVLSHQICEPHGCGSGSVVFCPLLEIPGCQDLAGEDAPAI 306
```

```
Db 239 CIGARVQAQSPNSGLDFQDNQVSVKLVPPKCHGSVIAF----- 278
Qy 307 QOSPLLCSSVRPEAGTSLSESWLLYSHPTNKKRRVDLGIYLNQSPLEAACWSRPWILHC 366
Db 279 -----PNPTSKADAL-DWLLYIUTTSRKTNTLGVNLQKPLDPTTWSAPTLLAT 328
Qy 367 GPCGYSDLAAL-----ENGLFGCLFECGCKQCEQIAFRFLTDRILSHVOG 414
Db 329 GICAYSDLQNHGHDGSPQSGCLYE---SNNYEIVELMFTLKOAFPAVFG 377

RESULT 7
US-09-791-537-87692
; Sequence 87692, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87692
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-87692

Query Match 10.9%; Score 255.5; DB 5; Length 415;
Best Local Similarity 28.0%; Pred. No. 2.9e-14;
Matches 116; Conservative 46; Mismatches 130; Indels 123; Gaps 23;

Qy 23 TYRIPALIVPPAHTFLAFAEKRS--SSKDEDALHLVLRRLTQSGVOWEPLKSLMK-AT 80
Db 76 TFRIP-LITATPRGTLAFAEARKMSSDEGAKFIALRRSMDQGST--WSPATFIVNDGD 132
Qy 81 LPHGHTMNPWPWKRKSGVYLFF-ICVOGHVTERQQIMSGRNPARLCFCISQDAGYSWS 139
Db 133 VPDGLNLG-AVVSDETGVVFLFYSLCAH-----KAGCOVASTMLVWSKDDGVWS 182
Qy 140 DVRLTEVIGPEVTHWATFVAGVGHGIGLO-----SGLIIPAYAYIIPFFWFCFLPYR 195
Db 183 TPRNLSD-IGTEV-----FAPGPGSGIQKQREPRKGLIIVCGHGTLEDRDGVFC 230
Qy 196 ARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVLYCSARTPNRRAEAL 255
Db 231 -----LLSDDHGASWRYGS-----GVSGIP----- 250
Qy 256 SIDHGEQFKVPLSHQICEPPHGCQGSV-----SFCPLEIPGGCQDLAGEDAPAIQ 307
Db 251 ---YGQPKQENDFPDECOPYELPDGVSVINARNQNNYHCHCRIV-----LRSYDA--- 298
Qy 308 QSPILCSSVRP-----BPE-----AGTL--SESWLLYSHPTNKKRRVDLGIYLNQSP 353
Db 299 -----CDTLRPRDVTDFPELDPVVAAGAVVTSSGIVFFSNPAHPFRVNL--TLRWSFS 351
Qy 354 EAACWSRPWI-LHCGPCGYSDLAALENEGLFG-----CLFECGCKQCEQIA 399
Db 352 NGTSWRKETVOLWPGSPGSSYSLATLEG-SMDGEQAQPLVLYVEKGRNHYTESIS 405

RESULT 8
US-10-217-623-464
; Sequence 464, Application US/10217623
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM031C1N
; CURRENT APPLICATION NUMBER: US/10/217,623
```

```
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/758,442
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-623-464

Query Match 10.9%; Score 255.5; DB 6; Length 425;
Best Local Similarity 28.0%; Pred. No. 2.9e-14;
Matches 116; Conservative 46; Mismatches 130; Indels 123; Gaps 23;

Qy 23 TYRIPALIVPPAHTFLAFAEKRS--SSKDEDALHLVLRRLTQSGVOWEPLKSLMK-AT 80
Db 86 TFRIP-LITATPRGTLAFAEARKMSSDEGAKFIALRRSMDQGST--WSPATFIVNDGD 142
Qy 81 LPHGHTMNPWPWKRKSGVYLFF-ICVOGHVTERQQIMSGRNPARLCFCISQDAGYSWS 139
Db 143 VPDGLNLG-AVVSDETGVVFLFYSLCAH-----KAGCOVASTMLVWSKDDGVWS 192
Qy 140 DVRLTEVIGPEVTHWATFVAGVGHGIGLO-----SGLIIPAYAYIIPFFWFCFLPYR 195
Db 193 TPRNLSD-IGTEV-----FAPGPGSGIQKQREPRKGLIIVCGHGTLEDRDGVFC 240
Qy 196 ARPHSLMIYSDDLGATWHHGRLLIKPMVTVECEVAEVIKAGHPVLYCSARTPNRRAEAL 255
Db 241 -----LLSDDHGASWRYGS-----GVSGIP----- 260
Qy 256 SIDHGEQFKVPLSHQICEPPHGCQGSV-----SFCPLEIPGGCQDLAGEDAPAIQ 307
Db 261 ---YGQPKQENDFPDECOPYELPDGVSVINARNQNNYHCHCRIV-----LRSYDA--- 308
Qy 308 QSPILCSSVRP-----BPE-----AGTL--SESWLLYSHPTNKKRRVDLGIYLNQSP 353
Db 309 -----CDTLRPRDVTDFPELDPVVAAGAVVTSSGIVFFSNPAHPFRVNL--TLRWSFS 361
Qy 354 EAACWSRPWI-LHCGPCGYSDLAALENEGLFG-----CLFECGCKQCEQIA 399
Db 362 NGTSWRKETVOLWPGSPGSSYSLATLEG-SMDGEQAQPLVLYVEKGRNHYTESIS 415

RESULT 9
US-09-791-537-1830
; Sequence 1830, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1830
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-1830

Query Match 10.5%; Score 245.5; DB 5; Length 409;
Best Local Similarity 27.1%; Pred. No. 2.2e-13;
Matches 112; Conservative 46; Mismatches 136; Indels 119; Gaps 22;
```


Qy 23 TYRIPALIVPPAHTFLAE--KRSSKDEDAHLVLRGLRTGQSVQWEPLKSLMKATL 81
Db 70 TFRIP-LITATPRGTLFAFAEARKKSDEGAKFIAMRS--TDGSSWSTAFIVDDGE 126
Qy 82 FGRMTNCPVWERKSGVYLFF-TCVQGHVTERQIIMSGRNPALCFCTCSQDAGYSWD 140
Db 127 ASDGLNLGAVNDVTGIVFLYITLC--AHKVNQC-----VASTMLVWSKDDGISNP 177
Qy 141 VRDLTEEVIGPEVTHWATFVCGHGILQ-----SGRLIIPAYAYIIPFWFCFLPYRA 196
Db 178 PRNLSVD-IGTEM-----FAPGSGGIQKQREPGKGRIVCGHGLERDGVFC----- 224
Qy 197 RPHSLMIYSDDLGATWHHGRKIPMTVECEVAEIVGKAGHPVLYCSARTPNHRAEALS 256
Db 225 -----LLSDDHGASWHVG-----TGVSIGP----- 244
Qy 257 IDHGEFCQKPVLSHOLCEPHPGHCGSVVFCPLFPGGCQDLAGED-----APAIQOS 309
Db 245 -----FGQPKHDHDF--NPDEQ-----PYELPDGSVIINARNQNNYHCRIVLRS 289
Qy 310 PLLCSSVRP-----EPE-----ACTL--SESWLLYSHPTNKKRRVDLGIYLNOSPLEA 355
Db 290 YDADCTLRPRDVTFDPELVDVPVVAAGALATSGGIFFSNPAHPEFRVNL--TLRWSFSNG 347
Qy 356 ACWSRPWI-LHCGPCGYSDLAALENE-----GLFGCLFECGKQCEQIA 399
Db 348 TSWQRQVQVWPGSGYSLSLTALENTGCKQPPQLF-VLYEKGLNRYTESIS 399
RESULT 10
US-09-791-537-73212
; Sequence 73212, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 73212
; LENGTH: 605
; TYPE: PRT
; ORGANISM: pdb 1EUT
US-09-791-537-73212

Query Match 8.6% Score 201.5; DB 5; Length 605;
Best Local Similarity 24.0%; Pred. No. 3.1e-09;
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;
Qy 12 PLFQOED-----KRGV-TYRIPALIVPPAHTFLAEKRSSSKDEDAHLVLR-RRGLRT 64
Db 7 PLYTEQDLAVNCREGFPNRYIPALT-VTPDGLLASYGDRPTGIDAPGNLSILQRRSTDG 65
Qy 65 GQSVQWEPLKSLMKATLPCHRTMPCPVWERKSGVYVLFICVQGHVTERQIIMSGRNP 124
Db 66 GRTWGEQVVSAGQTTAPIKGFSDPSYLVDRGTGTIFNF-----HVYSQRQGFAGSRPG 119
Qy 125 -----RLCFICSQDAGYSWSDVRLTTEVIGPEVTHWATFVCGHGILQ-----S 171
Db 120 TDPADPNVLHANVATSTGGTLTWSH-RTITAD-ITPD-PGWRSRFAASGEGIQLRYGPHA 176
Qy 172 GRL-----IIPAYAYIIPFWFCFLPYRAPHSLMIYSDDLGATWHHGRLI-----KPMV 222
Db 177 GRLLIOQYTIINAAGAF-----QAVSVYSDDHGRWTRAGEAVGVGMDEKN 220
Qy 223 TVECEVAEIVGKAGHPVLYCSARTPNHRAEALSIDHGEFCQKPVLSHOLCEPHPGHCGS 282

Db 221 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGCHSYGPVTTIDRLDPDPTN--NAS 271
Qy 283 VVSFCEPLEIPGCGQDLAGEDAPAIQOSPLLCSSVSRPEPEAGTLSWSWLLYSHPTNKKRRV 342
Db 272 IIRAPP-----DAP-----AGSARAKVLLFNSAASQTSR- 300
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGPCGYSDLAALENEGLFGCLFECGTKQCEQ 397
Db 301 -----SOGTIRMSCDDGQTPVSKVPFGSGMSYSTLTALP-DGYTGLLYEPGT-----G 348
Qy 398 IAFRLFTDREILSHVQDC-----STPGMNSPESKK 428
Db 349 IRYANFN-----LAWLGGICAPETIPDVALEPCQQ 378
RESULT 11
US-09-791-537-81425
; Sequence 81425, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 81425
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Micromonospora viridifaciens
US-09-791-537-81425
Query Match 8.6% Score 201.5; DB 5; Length 647;
Best Local Similarity 24.0%; Pred. No. 3.4e-09;
Matches 109; Conservative 65; Mismatches 161; Indels 119; Gaps 24;
Qy 12 PLFQOED-----KRGV-TYRIPALIVPPAHTFLAEKRSSSKDEDAHLVLR-RRGLRT 64
Db 49 PLYTEQDLAVNCREGFPNRYIPALT-VTPDGLLASYGDRPTGIDAPGNLSILQRRSTDG 107
Qy 65 GQSVQWEPLKSLMKATLPCHRTMPCPVWERKSGVYVLFICVQGHVTERQIIMSGRNP 124
Db 108 GRTWGEQVVSAGQTTAPIKGFSDPSYLVDRGTGTIFNF-----HVYSQRQGFAGSRPG 161
Qy 125 -----RLCFICSQDAGYSWSDVRLTTEVIGPEVTHWATFVCGHGILQ-----S 171
Db 162 TDPADPNVLHANVATSTGGTLTWSH-RTITAD-ITPD-PGWRSRFAASGEGIQLRYGPHA 218
Qy 172 GRL-----IIPAYAYIIPFWFCFLPYRAPHSLMIYSDDLGATWHHGRLI-----KPMV 222
Db 219 GRLLIOQYTIINAAGAF-----QAVSVYSDDHGRWTRAGEAVGVGMDEKN 262
Qy 223 TVECEVAEIVGKAGHPVLYCSARTPNHRAEALSIDHGEFCQKPVLSHOLCEPHPGHCGS 282
Db 263 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGCHSYGPVTTIDRLDPDPTN--NAS 313
Qy 283 VVSFCEPLEIPGCGQDLAGEDAPAIQOSPLLCSSVSRPEPEAGTLSWSWLLYSHPTNKKRRV 342
Db 314 IIRAPP-----DAP-----AGSARAKVLLFNSAASQTSR- 342
Qy 343 DLGIYLNOSPLEAAC-----WSRPWILHCGPCGYSDLAALENEGLFGCLFECGTKQCEQ 397
Db 343 -----SOGTIRMSCDDGQTPVSKVPFGSGMSYSTLTALP-DGYTGLLYEPGT-----G 390
Qy 398 IAFRLFTDREILSHVQDC-----STPGMNSPESKK 428
Db 391 IRYANFN-----LAWLGGICAPETIPDVALEPCQQ 420

```

RESULT 12
US-09-791-537-73210
; Sequence 73210, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73210
; LENGTH: 365
; TYPE: PRT
; ORGANISM: pdb leuR
US-09-791-537-73210

Query Match      8.5%; Score 200.5; DB 5; Length 365;
Best Local Similarity 24.4%; Pred. No. 2e-09;
Matches 101; Conservative 58; Mismatches 148; Indels 107; Gaps 21;

QY 12 PLFQQED-----KRGV--TYRIPALIIVPPAHTLAFIAEKRRSSSKDEDAHLVL--RRGLRT 64
   || :|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 7  PLTEQDLAVNGREGPNYRIPALT-VTPDGDLLASYDGPRPTGIDAPGPSILQRSTDG 65
   || :|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 GQSVMQEPLKSLMKATLPGHRTWNPFCWPWERKSGYVYLFFICVQGVHATERQQITMSGRNPA 124
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 66 GRWGEOQVVYSAGOTTAPIKGFSDPYSLVDRETGTINF-----HVYSORQGFAGSRPG 119
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 -----RLCFCSODAGYSNSDVRLTVEEVIGPETHWATFAVGPCHGIOLO----S 171
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 120 TDPADPNVLHANVATSTDGGLTWSH--RTITAD--ITPD-PGWRSRFAASGEGIOLRYGPHA 176
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 172 GLR-----IIPAYAYIYPFEFFCFLPYRAPRHPSLMITYSDDLGATWHHGRLI----KPMV 222
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 177 GLRIQYTIINAAGF-----QAQSVSYDDHGTWRAGEAVGVGMDENK 220
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 223 TVECEVAEVIKAGHPVLYGSARTPNHRHAELSIDHGECFQKPVLSHQLCEPHGCQGS 282
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 221 TVELSDGRVLLNSRD-----SAR--SGYRKVAVSTDGGHSYGVPVTIDRLDPDTN--NAS 271
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 283 VVSFCEPLEIPGCCODLAGEDAPAIOQSPLLCSSVRPEPEAGTISESWLLYSHPTNKKRV 342
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 272 IIRAFP-----DAP-----AGSARAKVLLFSNAASQTSR- 300
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 343 DLGIYNOSPLEAAC-----WSRPWILLHCPCGYSDLAALENEGFLCGFECGT 391
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 -----SQGTIRMSCDDGQTWPVKVFQPGQSMYSTLTALP-DGTGYLLYPEGT 347
   || :|| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-791-537-48863
; Sequence 48863, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48863
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-09-791-537-48863

```

Db 723 Y-COLSVINY-----SQIDCKDA-----IIFAIPDAN-----YNRV 754
Qy 337 NKRRVDL-----GYLNOSPLEAACWSRPWTLHCGPCGYSDLAALENEGLPGLFE-CGT 391
Db 755 NGTVRVGLITENGSENGEPRYDIEMRYNKVAPGTGYSCLSMPN-GEIGLFYEGRGS 813
Qy 392 KQ 393
Db 814 RQ 815

Search completed: October 7, 2002, 15:05:57
Job time: 439 sec

RESULT 15
US-09-791-537-140247
; Sequence 140247, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 140247
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Clostridium perfringens
US-09-791-537-140247

Query Match 7.3%; Score 172; DB 5; Length 694;
Best Local Similarity 20.8%; Pred. No. 1.6e-06;
Matches 102; Conservative 65; Mismatches 143; Indels 180; Gaps 25;

Qy 24 YRIPALIVVPPAHTFLAFAEKRSSKD--EDALHLVLR--GLRT--GOSVQWEPLK- 74
Db 265 YRIPALFKTREGTLIASIDARRHGADAPNNDIDTAVRRSEGGKTWDEGQIIMDYDPS 324
Qy 75 SLMKATLPGHRTMPCPWKRKSGYVYLF-----FICV- 107
Db 325 SVIDTTL-----IQDDETGRIFLLVTHPEPSKYGFNAGLGSFGKNDGKEYLCLY 374
Qy 108 -----QGHVTER-----QOIMSGRNPAP-----LCFI 129
Db 375 DSSGKEFTVRENVVYDKNGKTEYTNALGDLFRKNGTKIDINSTAPLKAKGTSYINLV 434
Qy 130 CSODAGYSWSDVRDLTERVIGPEVTHWATF-AVGPGHCIOLOS---CRLIIPAYAYIP 184
Db 435 YSDDDGKTWSEPNINQV-----KKDMWKFGLIAPGRGIIQKNGEHKGRIVVPVY---- 485
Qy 185 FWFEFCRLPY---RAPHSLMIYSDDLGAIW-----HHGRLIK-----PM 221
Db 486 -----YTNEKGKOSAVIYSDSGKNWTIGESPNDNRKLENGKIINSKTLSDDAPO 536
Qy 222 VTVECEVAETGKAGHPVLYCSARTPNHRAEALSIDHGECFQRPVLSHQCEPPHGCQG 281
Db 537 LT-ECQVEM--PNGQLKLF--MRNLGYLNIATSFDDGATWDETVEKDTNVLEPY-CQL 590
Qy 282 SVWSFCPLIEPGGCDLAGEDAPAIQOSPILCSSVRPEPEAGTLESWLLYSHPTNKKR- 340
Db 591 SVINY-----SOKVDGKA-----VIFSNPNARSRS 616
Qy 341 ----RVDL-----GIYLNOSPLEAACWSRPWTLHCGPCGYSDLAALENEGLPGLFECGTK 392
Db 617 NGTVRIGLINQVGYENGEPYEFDWKYNKLVKPGYAYSCLTSLN-GNIGLLYEGTPS 675
Qy 393 QCEQIAFRL 402
Db 676 EEMSVIEMNL 685

